



POPULATION ECOLOGY

Second
Edition

FIRST PRINCIPLES



John H. Vandermeer and Deborah E. Goldberg

VISIT...

LANZAROTE
Caliente.COM



Population Ecology



FIRST PRINCIPLES

Second Edition

JOHN H. VANDERMEER

DEBORAH E. GOLDBERG

PRINCETON UNIVERSITY PRESS

Princeton and Oxford

Copyright © 2013 by Princeton University Press

Published by Princeton University Press, 41 William Street, Princeton, New Jersey 08540

In the United Kingdom: Princeton University Press, 6 Oxford Street, Woodstock, Oxfordshire
OX20 1TW

press.princeton.edu

All Rights Reserved

ISBN 978-0-691-16030-6

ISBN (pbk.) 978-0-691-16031-3

Library of Congress Control Number: 2013940462

British Library Cataloging-in-Publication Data is available

This book has been composed in Sabon LT Std with Gotham Narrow display by
Princeton Editorial Associates Inc., Scottsdale, Arizona.

Printed on acid-free paper ∞

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Dedicated to the memory
of our colleague and dear friend
Beverly Rathcke

CONTENTS

List of Figures xi

List of Tables xvii

Preface xix

ONE **Elementary Population Dynamics** 1

Density Independence: The Exponential Equation 2

Density Dependence 9

The Logistic Equation 13

The Yield–Density Relationship 17

Density Dependence and Mortality: Thinning Laws 22

Density Dependence in Discrete Time Models 28

TWO **Projection Matrices: Structured Models** 30

Elementary Age-Structured Population Projection Matrices 30

Non-Age Structure: Stage Projection Matrices 39

Eigenvectors, Reproductive Value, Sensitivity, and Elasticity 45

Density Dependence in Structured Populations 48

Density Dependence in a Simple Age-Structured Model 48

Density Dependence in Size-Distributed Populations 50

Density Dependence in a Stage-Structured Model 56

Appendix: Basic Matrix Manipulations 57

	Matrix Multiplication	57
	Matrix Addition and Subtraction	58
	The Identity Matrix	59
	The Determinant of a Matrix	59
THREE	Applications of Simple Population Models	62
	Life History Analysis	63
	<i>Investment in Survivorship versus Reproduction: The r-K Continuum</i>	64
	<i>The Cost of Reproduction</i>	66
	<i>Optimal Reproductive Schedules</i>	67
	Applications of Population Projection Matrices	73
	<i>The Dall's Mountain Sheep: A Static Life Table</i>	73
	<i>Palo de Mayo: A Dynamic Life Table</i>	74
	<i>Population Viability Analysis</i>	76
	<i>Demography of Invasive and Native Plant Populations</i>	78
FOUR	A Closer Look at the "Dynamics" in Population Dynamics	81
	Intuitive Ideas of Equilibrium and Stability	83
	Eigenvalues: A Key Concept in Dynamic Analysis	92
	Basic Concepts of Equilibrium and Stability in	
	One-Dimensional Maps	97
	<i>The One-Dimensional Map</i>	98
	<i>Stability and Equilibrium in the Logistic Map</i>	106
	<i>Basins of Attraction in the Logistic Map</i>	108
	Structural Stability	110
	Bifurcation Diagrams	116
	Concluding Remarks	122
FIVE	Patterns and Dynamics in Space	126
	The Poisson Distribution	129
	Point Pattern Analysis and the Question of Scale	134
	Mechanisms of Spatial Pattern Formation:	
	Principles of Reaction/Diffusion	137
	Mechanisms of Spatial Pattern Formation: Biological Causes	141

	Metapopulations	142
	<i>Assumptions of Metapopulation Models</i>	146
	<i>The Rescue Effect and Propagule Rain</i>	148
	Appendix: Data for Exercises 5.2, 5.3, and 5.4	150
SIX	Predator–Prey (Consumer–Resource) Interactions	152
	Predator–Prey Interactions: First Principles	153
	Density Dependence	158
	Functional Response	161
	Functional Response and Density Dependence Together	166
	Paradoxes in Applications of Predator–Prey Theory	168
	Predator–Prey Dynamics: A Graphical Approach	170
	Predator–Prey Interactions in Discrete Time	176
SEVEN	Disease Ecology	187
	Direct Disease Transmission	188
	Indirect Transmission	194
EIGHT	Competition	198
	Competition: First Principles	199
	Isocline Analysis of the Lotka–Volterra Competition Equations	203
	Niches and Competitive Coexistence and Exclusion	209
	The Competitive Production Principle: Applications of Competition Theory to Agriculture	211
	Resource Competition	212
NINE	Facilitation and Mutualism	225
TEN	What This Book Was About	239
	<i>Glossary</i>	243
	<i>References</i>	247
	<i>Index</i>	255

FIGURES

- 1.1 Graphs of equation 5 5
- 1.2 Plot of aphid data (from table 1.1) 7
- 1.3 Aphid data from figure 1.2 plotted arithmetically 8
- 1.4 Growth of a culture of *Paramecium bursaria* in a test tube 13
- 1.5 Logarithmic plot of the data from figure 1.4 14
- 1.6 Long-term data for the same populations of *Paramecium bursaria* 15
- 1.7 Fit of logistic equation to the *Paramecium bursaria* data from figure 1.6 17
- 1.8 Exemplary yield versus density data for maize, rape, and beets 18
- 1.9 Equation of Shinozaki and Kira superimposed on the rape data from figure 1.8 19
- 1.10 Diagrammatic representation of the process of thinning 23
- 1.11 Typical relationship between log density of surviving plants and log dry weight 24
- 1.12 Expected pattern of growth and mortality as the thinning process and growth in biomass interact 25
- 1.13 Relationship between the density of genets and the mean weight per genet 26
- 1.14 Yoda and colleagues' interpretation of the origin of the three-halves thinning law 27
- 2.1 Population trajectories for different starting points in an age-distributed population 35

- 2.2 Long-term projections of the hypothetical tree population discussed in the text 43
- 2.3 Diagrammatic summary of the general structure of a stage projection matrix 45
- 2.4 Time series resulting from density dependence in a simple age-structured model 50
- 2.5 Size frequency distributions of two taxa of mangrove seedlings 51
- 2.6 Intraspecific competition and growth in populations of a limpet 53
- 2.7 An experiment on vines in which root and shoot competition were separated 54
- 3.1 Difference in per capita egg production between the O lines and B lines from Rose and Charlesworth's experiment 67
- 3.2 Alternative solutions for the Schaffer model of energy allocation 69
- 3.3 Elasticity of survival, growth, and fecundity for invasive and native species as a function of life span 79
- 4.1 Illustration of the dynamical behavior associated with a point attractor 84
- 4.2 Illustration of the dynamical behavior of a point repeller 84
- 4.3 Physical models of a classical attractor and repeller 85
- 4.4 Beaker and magnet model of dynamics 86
- 4.5 Traditional representations of an oscillatory attractor and an oscillatory repeller 87
- 4.6 Physical model illustrating a periodic attractor 88
- 4.7 Cross section (Poincaré section) through the surface of figure 4.6 88
- 4.8 Poincaré section with a strange attractor rather than a periodic attractor 89
- 4.9 State space for a one-dimensional (one-variable) model 92
- 4.10 The vectors of the example from figure 4.9 rotated 92
- 4.11 State space for a one-dimensional model based on the logistic equation 93
- 4.12 Dynamics of the logistic equation in one dimension 93
- 4.13 Physical model of the dynamics of a point attractor in two dimensions 95
- 4.14 Conditions of eigenvalues for the three most common qualitatively distinct arrangements in two dimensions 96
- 4.15 Conditions of eigenvalues for the two most common qualitatively distinct arrangements in two dimensions 96
- 4.16 Step-by-step illustration of the process of stair-stepping for a one-dimensional map 99

- 4.17 Step-by-step projection using a function rather than numerical values 100
- 4.18 Exponential equation presented as a one-dimensional graph 101
- 4.19 Graph of equation 3, illustrating a point attractor 101
- 4.20 Graph of equation 4, illustrating an unstable equilibrium 102
- 4.21 Graphs of equation 5, illustrating an oscillatory attractor and repeller 103
- 4.22 Eigenvalues for the various forms of stability in a one-dimensional map 105
- 4.23 Graphs of equation 5, illustrating a two-point periodic attractor, a strange attractor, and an oscillatory repeller 105
- 4.24 The difference between stability and instability in the regional sense 107
- 4.25 Graph of equation 6, illustrating the two basins of attraction for the two point attractors 109
- 4.26 Illustration of a structurally unstable parameter configuration for equation 16 111
- 4.27 Graphs of equation 8 (the logistic map), illustrating the structurally unstable configuration obtained when $\lambda = 2.0$ 112
- 4.28 Graphs of equation 8 (the logistic map), illustrating the structurally unstable configuration obtained when $\lambda = 3.0$ 113
- 4.29 Illustration of a saddle-node bifurcation 113
- 4.30 Example of a possible application of a saddle-node bifurcation 115
- 4.31 Illustration of a basin boundary collision 116
- 4.32 Illustration of the basic bifurcation process 117
- 4.33 Bifurcation diagram for the logistic map 118
- 4.34 Close-up of part of the bifurcation diagram of figure 4.33 119
- 4.35 Bifurcation diagram of the *Tribolium* model 120
- 4.36 Experimental results of a *Tribolium* experiment 121
- 4.37 Theoretical situations in which repellers and attractors coexist 123
- 5.1 Distribution of two species of trees on a 1-hectare plot 128
- 5.2 Birds on a telephone line 128
- 5.3 All possible ways that two birds can fit on three places on a telephone line 129
- 5.4 The telephone line divided into segments, each of which could contain a large number of birds 130
- 5.5 The graph in figure 5.1 with grid lines dividing the plot into 100 quadrats 131
- 5.6 Comparison of expected with observed density of two species of trees 132

5.7	The upper right quarters of the plots shown in figures 5.1 and 5.5	134
5.8	Stylized version of Ripley's K analysis, showing Ripley's K as a function of spatial scale	135
5.9	Spatial distribution of two tree species at the scale of 4 hectares	136
5.10	Spatial distribution of two tree species at the scale of 9 hectares	136
5.11	Artificial example of the diffusion of individual plants in a one-dimensional space	139
5.12	Spatial distribution of individual plants using equation 6	141
5.13	Examples of Turing patterns emerging from equations 7a and 7b	142
5.14	Diagrammatic representation of the Turing effect	143
5.15	Simplified classification of spatially structured dynamics	144
5.16	Results of seed addition experiments	147
5.17	Differences in extinction rates (the rescue effect) and immigration rates (the propagule rain)	149
6.1	Qualitative construction of isoclines separating regions of space where prey increases or decreases and predator either increases or decreases	155
6.2	Construction of trajectories from the combination of the prey and predator isoclines in figure 6.1	156
6.3	Examples of oscillating predator–prey dynamics	157
6.4	The changes in the isoclines and population trajectory from adding density dependence to the prey	160
6.5	Type II functional response of ladybird beetles at different densities of aphids	162
6.6	Forms of the functional response	163
6.7	The changes in the isoclines and population trajectory from adding a type II functional response	165
6.8	Changing dynamics due to adding both density dependence and a nonlinear functional response to the predator–prey equations	167
6.9	Trajectories of the predator–prey system	168
6.10	The paradox of enrichment	169
6.11	Isocline arrangements illustrating the paradox of biological control	170
6.12	The two dynamic functions that make up the prey dynamic system	171
6.13	Constructing the prey isocline from the attractors and repellers of the graph of rate versus density	173
6.14	Constructing the prey isocline using graphic arguments	174
6.15	The prey isocline and its dynamics under conditions of a prey refuge	175

- 6.16 The three forms of the functional response 176
- 6.17 Constructing the prey isocline from the attractors and repellers of the rate-versus-density graph 177
- 6.18 Results from the discrete form of the predator–prey equations 179
- 6.19 The probability of avoiding attack by the predator as a function of the net attack rate 181
- 6.20 Solution of equations 16a and 16b 182
- 6.21 Solution of equations 17 and 16b 183
- 6.22 Solution of equations 17 and 16b with different parameters 184
- 6.23 Dynamic trajectories of the equation set 17 and 16b 185
- 6.24 Phase plane representation of equation set 17 and 16b illustrating an invariant loop, a deformed invariant loop, a fragmented attractor, and a possibly chaotic attractor 185
- 7.1 Time course of an infection 189
- 7.2 The basic idea of the SIR model 191
- 7.3 Typical trajectories from the classical SIR model 192
- 7.4 Expanded view of figure 7.3B illustrating the meaning of the recovery/transmission ratio 193
- 7.5 Isoclines and vector fields for equation set 6 196
- 7.6 The two qualitatively distinct arrangements of the isoclines of equation system 6 196
- 8.1 Diagrammatic representation of the meaning of carrying capacity and competition coefficient 202
- 8.2 Isoclines of the competition equations 204
- 8.3 Putting both isoclines on the same graph and summing the vectors to get the vector field of the two-dimensional system 205
- 8.4 The four classic cases of interspecific competition 206
- 8.5 Physical model illustrating the behavior of a saddle point repeller 207
- 8.6 Explanation of the competitive exclusion principle 208
- 8.7 The graphic interpretation of the land equivalent ratio criterion 213
- 8.8 Relationship between birth and mortality rates and resource availability, and growth patterns of consumer and resource populations over time 217
- 8.9 Resource-dependent growth and mortality rates for two species using the same resource and the time course expected 218
- 8.10 Consumption and supply vectors and resource equilibrium for two essential resources 219
- 8.11 Zero net growth isoclines (ZNGIs) for substitutive and essential resources 220

8.12	The four distinct cases of resource competition	221
8.13	Equilibrial consumption and supply vectors in monocultures for two competitors with different supply points	223
9.1	The basic dynamics of equations 1a and 1b, showing the two fundamental outcomes of facilitative mutualism	227
9.2	The various forms of isocline placement in the basic mutualism model	228
9.3	Complications with obligate mutualisms	231
9.4	Summary of the qualitative results with symmetrical mutualisms	232
9.5	Transforming the isocline of equation 1a	233
9.6	Dynamic consequences of adding nonlinearity in the facilitative effects of each of the mutualists for high mutualism coefficients	234
9.7	The eight qualitatively distinct cases of mutualistic interactions	236
9.8	Production of latex as a function of the proportion of roots infected with mycorrhizal fungi	237
10.1	Diagrammatic representation of the subject matter of the text	239

TABLES

- 1.1 Number of aphids observed per plant in a field of corn and beans in Guatemala 7
- 3.1 Attributes of r - and K -selected species and the environmental characteristics that select for them 65
- 3.2 Static life table for Dall's mountain sheep 74
- 3.3 Results of a population viability analysis for Florida manatees under different hypothetical scenarios 77

PREFACE

When ecology emerged from the general subject of biology, it did so in the tradition of the world's great naturalists—Maria Sibylla Merian, Thomas H. Huxley, Alfred Russel Wallace, Charles Darwin, and others. When Ernst Haeckel finally coined the term “ecology,” the field was little more than what natural history had always been, the detailed study of the way organisms interact with one another and their environment, a series of fascinating stories about nature. Use of quantitative techniques was minimal. Larry Slobodkin, one of the pioneers in the development of theory, once quipped to a class that an ecologist is “someone who wears khaki pants, doesn't know much about mathematics, and misidentifies local beetles.”

Indeed, maturation of the field emerged from a more serious mathematical approach. The foundational equations of Lotka and Volterra were invented in 1926, and their application to a laboratory system by Gause was complete by 1935. Yet what turned out to be foundational was largely ignored until the late 1960s, when Robert MacArthur, Richard Levins, and Robert May began their illustrious careers, with many others to follow. Mathematically based theory became an acceptable, even essential ingredient, with increasingly sophisticated theoretical approaches.

As a consequence, the average ecology student today faces some daunting literature. The legacy of the late 1960s and early 1970s has been a fascinating, yet sometimes perplexing, collection of theoretical approaches that have transformed the field dramatically. A course in ecology today is as likely to contain complicated differential equations as descriptions of life histories, a dramatic change from the situation 40 years ago, when mathematics was little mentioned. And observational studies as much as experiments now rely on predictions from abstract theory. Ecology has become a remarkably exciting

discipline principally because of this burgeoning theoretical superstructure.

Along with the momentum derived from this paradigm shift, we also see some lingering frustration as students confront this literature. In our experience, this frustration is frequently derived from a less than adequate appreciation of what have become the fundamental quantitative principles of ecology. In this book we have attempted to present these fundamentals in as heuristic a way as possible.

But precisely what are those fundamentals? Attempting to summarize the underlying quantitative framework for the entire field of ecology can be frustrating. Indeed, the content of a course in “general ecology” at a U.S. university may vary considerably. We found the choice and arrangement of subject matter only loosely correlated when comparing recent texts. This variability is, we believe, the consequence of the youthfulness of our science. Although this may seem frustrating to some, it is a source of excitement for most in the field.

Nevertheless, there is one subdiscipline of ecology that seems to have developed a canon. Virtually all courses that contain *population ecology* in their titles cover essentially the same material. In this one corner of the field, ecologists seem to agree on what constitutes the basic subject matter. And any survey of contemporary literature in ecology will uncover one or more of the basic ideas of population ecology underlying almost every published study.

In this text we attempt to present these first principles of population ecology. The book is intended as a text for advanced undergraduate and beginning graduate students, and it focuses on the analytical details of the basic subject matter. It is not at all intended to be an introduction to the literature. Individual applications are chosen as examples of the models and techniques presented and are not intended to be the best or the best-known studies in the literature—they are simply examples. Indeed, a review of the literature of “population ecology” would fill several volumes and would likely not serve much purpose anyway. In this book we have not made such an effort.

Although the subject matter of population ecology and thus this book is highly mathematical, the material is presented in such a way that only introductory calculus and a basic understanding of linear algebra are necessary prerequisites. For those students far removed from their calculus class, a basic understanding of the nature of a derivative and an integral plus the ability to differentiate simple functions (polynomials and exponentials) is really all that is required to fully appreciate this basic material. On the other hand, many students of ecology do not come to the field with a proper background in the basic operations of linear algebra. For these students we include an appendix at the end of Chapter 2.

This edition of the book includes an extensive set of exercises peppered throughout the text. These exercises are intended to encourage the student to think about the development of the text *before* that development happens. The exercises range from trivially simple to frustratingly complex. Our experience is that the attempt to solve a problem is the pedagogically important part of an exercise. We encourage students to develop their own style of deal-

ing with the exercises, ranging along a continuum. At one end, simply ignore the exercises. The text is, like the previous edition, intended as a complete document without the exercises anyway, so students who wish to receive only a superficial introduction or students who feel especially confident with the material and wish only to have their understanding of the subject refreshed can easily skip all of the exercises. At the other extreme of the continuum, it is, we feel, worthwhile to attempt to do each of the exercises before reading on. As in the case of any quantitative approach, the exercises may sometimes be frustrating, especially if one begins thinking about them in the wrong direction. We thus recommend that when an exercise begins to seem intractable or when frustration begins, even at the initial formation of the answer, the student move to the answer sheet (which is supplied at <http://www.sitemaker.umich.edu/jvander/home> in the form of an editable spreadsheet). Most of the answers are evident just from examining the worksheet. But it may be more useful for some students that the formulas in the cells of the spreadsheet can be easily manipulated by anyone with a knowledge of basic spreadsheet operation. The ability for such manipulation provides the student with the option of experimenting with various parameter values and starting conditions from which to empirically explore the subject matter of the exercise.

The organization of the book is based on chapter 1 as a foundation. With the exception of chapters 8 and 9, which depend to some extent on material from chapter 6, the rest of the chapters could stand alone after chapter 1 has been read. Chapter 2 enlivens the subject matter of chapter 1 with the reality that most populations in nature have some sort of structure. Further elaborations of this introductory material can be found in Caswell's excellent text (Caswell 2001). Chapter 3 follows with some simple applications of chapters 1 and 2 to some theoretical and practical issues. Chapter 4 introduces to the basic models nonlinearities—a subject that has witnessed an explosion in the literature—although most of the treatments are rough going mathematically. References that should be accessible to most ecology students are the two-volume set by Jackson (1991), the less inclusive but perhaps simpler text by Alligood et al. (1996), and, perhaps the most enlightening, the general text by Strogatz (2001). Chapter 5 deals both with statistical descriptions of spatial aggregation of individuals and populations and with population-dynamic models that incorporate spatial information either as foundations for the dynamics (e.g., metapopulations) or as consequences of the dynamics (reaction/diffusion equations). Spatial statistics is a burgeoning field, especially with the explosive growth in the use of geographic information systems. Our treatment is elementary, and students interested in this subject need to consult more sophisticated treatments (e.g., Diggle 1983; Fortin and Dale 2005). Metapopulation theory is introduced in its traditional form strictly as a mean field theory, and spatial pattern formation uses a simple qualitative presentation of reaction/diffusion equations. More detailed elaborations of the metapopulation concept can be found in the monograph by Hanski (1999), and the formation of spatial pattern is explored in several chapters of the volume edited by Dieckmann et al. (2000).

Finally, chapters 6–9 are introductions to two-species interactions. Chapter 6 deals with positive–negative interactions generally, specifically formulated as predator–prey interactions. More sophisticated treatments of this material can be found in the many applications of predator–prey theory in the literature, but we know of no particular text that treats advanced predator–prey theory as such. Useful applications can be found in Hawkins and Cornell (1999). Epidemiology is really the conceptual application of predator–prey theory to microparasites, but the framework that is used is basically that of metapopulations, as discussed in chapter 7. The standard reference on this material has become Anderson and May (1991), and a more recent and more compact summary can be found in Keeling and Rohani (2008). Chapter 8 deals with interspecific competition. More complex elaborations of the material presented in this chapter can be found elsewhere (e.g., MacArthur 1972; Tilman 1982; Chesson 2000). Finally, chapter 9 briefly outlines the fundamental theoretical ideas of facilitation and mutualism.

All of the material in this book has been presented to several groups of graduate and advanced undergraduate students at the University of Michigan. We owe a debt of gratitude to those students for their contributions to improvements that led to the present text. We also wish to thank our colleagues Mark Wilson, who contributed substantially to initial drafts, as well as Mercedes Pascual, Annette Ostling, and Aaron King, who were kind enough to use it in their classes. Special thanks to Dave Allen, Doug Jackson, Ed Baskerville, and Gyorgy Barabas for pointing out errors and better ways of presenting things over the years.



Elementary Population Dynamics

1

In 1960 the famous cyberneticist Heinz von Foerster and colleagues devised an equation predicting that the human population would become effectively infinite on Friday the 13th of November, 2026, meaning that at that point in time all humans would perish because the next individual to be born would crush everyone else—mass death due to squashing! In fact von Foerster and his colleagues were making a tongue-in-cheek argument to call attention to an issue they thought quite important. This was one, perhaps humorous, example of the application of simple quantitative principles of population dynamics to problems considered important.

Indeed there are many contexts in which it is important to understand the quantitative characteristics of single populations of organisms. In fisheries management, for example, the manager is interested in being able to predict the density of a fish population in the future under different management plans. An agronomist may wish to know the yield of a population of maize plants when planted at a particular density; an epidemiologist will want to know the density of disease-infected humans next month. Many other examples could be cited with clear practical importance. Of perhaps even more importance are theoretical applications that give us a more detailed understanding of more complex ecological systems. We might be interested in knowing the rate at which a population changes its density in response to selection pressure as part of a general program of understanding the consequences of natural selection under some hypothetical or real constraints. These topics, both applied and theoretical, are typical of the field called population ecology, and they all start with some basic ideas of what single populations of organisms do.

The unit of analysis is, not surprisingly, the “population,” a concept that is at once simple and complicated. The simple idea is that a population is simply a collection of individuals. But, as most ecologists intuitively know, the idea of a population is considerably more complex when one deals with it in any of

the applied or theoretical contexts alluded to above. To know what size limits one should place on a fish species one must know not only the number of fish in the population but the size distribution of that population and how that distribution relates to the population's reproductive effort. To decide when to take action on the emergence of pest species in an agricultural or forestry context, the distribution of individuals in various life stages must be known. In deciding whether a species is threatened with extinction, its distribution in space and movement among subpopulations (i.e., metapopulation dynamics) is far more important than simply its numerical abundance. And, to use the most frequently cited example, given the huge variation in the consumption of resources per person around the globe, the absolute numbers of the human population may be much less important than the activities undertaken by the members of that population—doomsday could be at hand well before 2026.

Thus the subject of population ecology can be a very complicated one indeed. But, as in the case of any science, we begin by assuming that it is rather simple. We eliminate the complications, make simplifying assumptions, and try, as much as possible, to develop general principles that might form a skeleton onto which the flesh of real-world complications might meaningfully be attached. In 2013 it seems that, unlike some other fields of biology, population ecology has a certain core subject matter that has come to be the “conventional wisdom.” That is, wherever a university course in population ecology is taught, pretty much the same material is covered, at least at the beginning of the course. This text is our attempt to present that core material precisely, and this chapter covers the first two essential ideas—the density independence and density dependence of population growth.

Density Independence: The Exponential Equation

It is often surprising how quickly a self-reproducing event becomes a big event. The classic story is this: suppose you have a pond with some lily pads in it, and suppose each lily pad replicates itself once per week. If it takes a year for half the pond to become covered with lily pads, how long will it take for the entire pond to become covered? If one does not think too long or too deeply about the question, the quick answer seems to be about another year. But with a moment's reflection one can retrieve the correct answer—only one more week.

This simple example has many parallels in real-world ecosystems. A pest building up in a field may not seem to be a problem until it is too late. A disease may seem much less problematical than it really is. The simple problem of computing the “action threshold” (the population density a pest must reach before you have to spray pesticide) requires the ability to predict how large a population will be based on its prior behavior. If half the plants in the field are attacked within three months, how long will it be before they are all attacked?

To understand even the extremely simple example of the lily pads, one constructs a mathematical model, usually quite informally, in one's head. If

all the lily pads on a pond replicate themselves once per week, in a pond half-filled with lily pads each one of those lily pads will replicate itself in the next week and thus the pond will be completely filled up. To make the solution to the problem general, we simply say the same thing, but instead of labeling the entities lily pads, we call them something general, say organisms. If organisms replicate themselves once per week, by the time the environment is half full, it will take only one week to become completely full. Implicitly, the person who makes such a statement is simply stating verbally the following equation:

$$N_{t+1} = 2N_t. \quad (1)$$

N is the number of organisms, in this case lily pads. Instead of N , say lily pads, and instead of t , say this week, and instead of $t + 1$, say next week, and equation 1 expresses simply, “The number of lily pads next week will be equal to twice the number this week.”

Writing down equation 1 is no different than making any of the statements that were made previously about it. But by making it explicitly a mathematical expression, we bring to our potential use all the machinery of formal mathematics. And that is actually good, even though beginning students sometimes don’t think so.

Using equation 1 we can develop a series of numbers that reflect the changes of population numbers over time. For example, consider a population of herbaceous insects. If each individual produces a single offspring once per week, and if those offspring mature and each produces an offspring within a week, we can apply equation 1 to see exactly how many individuals will be in the population at any point in time. Beginning with a single individual, we have, in subsequent weeks, 2, 4, 8, 16, 32, 64, 128, and so on. If we change the conditions such that the species replicates itself twice per week, equation 1 becomes

$$N_{t+1} = 3N_t \quad (2)$$

(with a 3 instead of a 2, because before we had the individual and the single offspring it produced but now we have the individual and the two offspring it produced). Now, beginning with a single individual, we have, in subsequent weeks, 3, 9, 27, 81, 243, and so on.

We can use this model in a more general sense to describe the growth of a population for any rate of production of offspring at all (not just 2 and 3 as above). That is, write

$$N_{t+1} = \lambda N_t \quad (3)$$

where λ can take on any value at all. λ is frequently called the finite rate of population growth (or the discrete rate).

It may have escaped notice in the above examples, but either of the series of numbers could have been written with a much simpler mathematical notation. For example, the series 2, 4, 8, 16, 32, is actually $2^1, 2^2, 2^3, 2^4, 2^5$, while the series 3, 9, 27, 81, 243 is actually $3^1, 3^2, 3^3, 3^4, 3^5$. So we could write

$$N_t = \lambda^t, \quad (4)$$

which is just another way of representing the facts described by equation 3. (Remember, we began with a single individual, so $N_0 = 1.0$.)

EXERCISES

- 1.1 Compute the expected number of individuals over time in a population growing according to equation 3 for $\lambda = 1.9$, $\lambda = 2.0$, and $\lambda = 2.1$, where λ is the finite rate of increase (it is convenient to begin with a single individual). Graph the results as a function of time.
- 1.2 Plot the numbers calculated in exercise 1.1 as $\ln(N)$ versus time. Sketch (by hand, by eye) the function that best approximates one of the data sets (try for $\lambda = 2$), and visually estimate what the parameters of that function might be. What is the nature of the function, and how do its parameters relate to the original construction of the set of numbers in exercise 1.1?

For further exposition we wish to express the constant λ in a different fashion. It is a general rule that any number can be written in a large number of ways. For example, the number 4 could be written as $8/2$ or $9 - 5$ or 2^2 or in many other ways. In a similar vein, an abstract number, say λ , could be written in any number of ways: $\lambda = 2^b$, where b is equal to $\lambda/2$, or $\lambda = 2^b$, in which case $b = \ln(\lambda)/\ln(2)$ (where $\ln$ stands for natural logarithm). For reasons that will be obvious to the reader not too far removed from the elementary calculus class, if we represent λ as 2.7183^r , a powerful set of mathematical tools immediately becomes available. The number 2.7183 is Euler's constant, usually symbolized as e (actually, 2.7183 is rounded off and thus is only approximate). It has the important mathematical property that its natural logarithm is equal to 1.0 or $\ln(e) = 1$.

So we can rewrite equation 4 as

$$N_t = e^{rt}, \quad (5)$$

which is the classical form of the exponential equation (where λ has been replaced with e^r). One more piece of mathematical manipulation is necessary to complete the toolbox necessary to model simple population growth. Another seemingly complicated but really rather simple relationship that is always learned (but frequently forgotten) in elementary calculus is that the rate of change of the log of any variable is equal to the derivative of that variable divided by the value of the variable. This rule is more compactly stated as

$$\frac{d(\ln N)}{dt} = \frac{dN}{Ndt}. \quad (6)$$

So take the natural logarithm of equation 5 as

$$\ln(Nt) = rt$$

and differentiate with respect to t to obtain

$$\frac{d(\ln N)}{dt} = r, \quad (7)$$

and we can then use equation 6 to substitute for the left-hand side of equation 7 to obtain

$$\frac{dN}{Ndt} = r.$$

After multiplying both sides by N , we obtain

$$\frac{dN}{dt} = rN. \quad (8)$$

Equations 5 and 8 are the basic equations that formally describe an exponential process. Equation 8 is the differentiated form of equation 5, and equation 5 is the integrated form of equation 8. They are thus basically the same equation (and indeed are quite equivalent to the discrete form—equation 3). Depending on the use to which they are to be put, any of the above forms may be used, and in the ecological literature one finds all of them. Their basic graphical form is illustrated in figure 1.1.

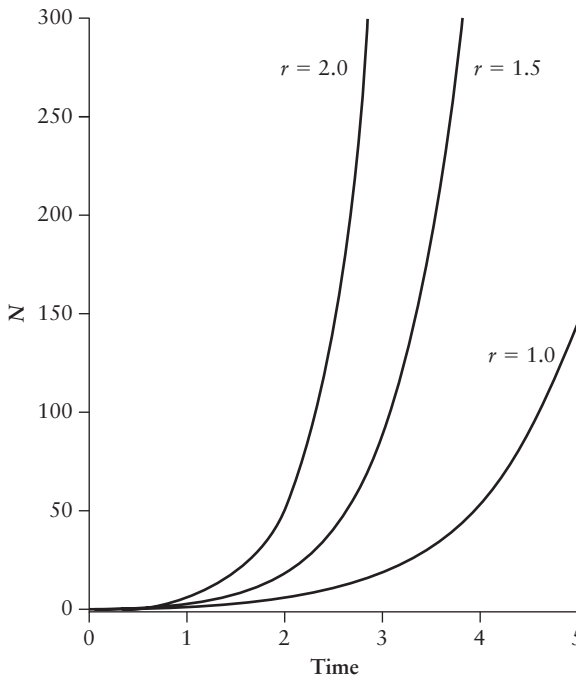


FIGURE 1.1. Graphs of equation 5.

EXERCISES

- 1.3 Derive an equation for doubling time in an exponentially growing population, and derive one for tripling time.
 - 1.4 Plot the values computed in exercise 1.1 on a graph of N_{t+1} versus N_t , where N is the number of individuals and t is time. What is the shape of the plotted points? What would be the value of a regression coefficient through the points? How does that relate to the original values of λ ?
 - 1.5 The human population in a small island nation is said to be doubling every 20 years. What would be the value of r (assuming an exponential model)?
 - 1.6 Using the differential form (i.e., continuous time), suppose that $r = 0.5$. Compute N as a function of t , again beginning with a single individual. Do the same for $r = 1.5$ and for $r = 1.0$. Graph all three results, and think about the pattern.
-

In the examples of exponential growth introduced above, the parameter r was introduced as a birth process only. The tacit assumption was made that there were no deaths in the population. In fact, all natural populations face the reality of mortality, and the parameter of the exponential equation is really a combination of birth and death rates. More precisely, if b is the birth rate (number of births per individual per time unit) and d is the death rate (number of deaths per individual per time unit), the parameter of the exponential equation is

$$r = b - d, \quad (9)$$

where the parameter r is usually referred to as the intrinsic rate of natural increase.

One other simplification was incorporated into all of the above examples. We always presumed that the population in question was initiated with a single individual, which almost never happens in the real world. But the basic integrated form of the exponential equation is easily modified to relax this simplifying assumption. That is,

$$N_t = N_0 e^{rt}, \quad (10)$$

which is the most common form of writing the exponential equation. Thus there are effectively two parameters in the exponential equation, the initial number of individuals, N_0 , and the intrinsic rate of natural increase, r .

Putting the exponential equation to use requires estimation of these two parameters. Consider, for example, the data presented in table 1.1. Here we have a series of observations over a five-week period of the number of aphids on an average corn plant in an imaginary cornfield.

As a first approximate assumption, let us assume that this population originates from an initial cohort that arrived in the field on March 18 (one week prior to the initial sampling). We can apply equation 10 to these data most easily by taking logarithms of both sides, thus obtaining

$$\ln(N_t) = \ln(N_0) + rt, \quad (11)$$

TABLE 1.1. Number of Aphids Observed per Plant in a Milpa (Corn and Beans) in the Highlands of Guatemala (Morales, 1998)

Date	Number of aphids	$\ln(\text{Number of aphids})$
March 25	0.02	-3.91
April 1	0.50	-0.69
April 8	1.50	0.40
April 15	5.00	1.61
April 22	14.50	2.67

which gives us a linear equation relating the natural logarithm of the number of aphids to time (where we code March 18 as time = 0, March 25 as time = 1, April 1 as time = 2, etc.). In figure 1.2 is a graph of this equation along with the original data points, and in figure 1.3 is a graph of the original data along with the fitted curve on arithmetic axes.

From these data we have the estimate of 1.547 aphids per parent aphid per week added to the population (i.e., the intrinsic rate of natural increase, r , is 1.547, which is the slope of the line in figure 1.2). The intercept of the regression is -4.626, which indicates that the initial population was .0098 (i.e., the antilog of -4.626 is .0098), which is an average of about one aphid per 100 plants. Now, if we presume that once the plants become infected with more than 40 aphids per plant the farmer must take some action to try to control them, we can use this model to predict when, approximately, this time will arrive. The regression equation is

$$\ln(N \text{ of aphids}) = -4.626 + 1.547t,$$

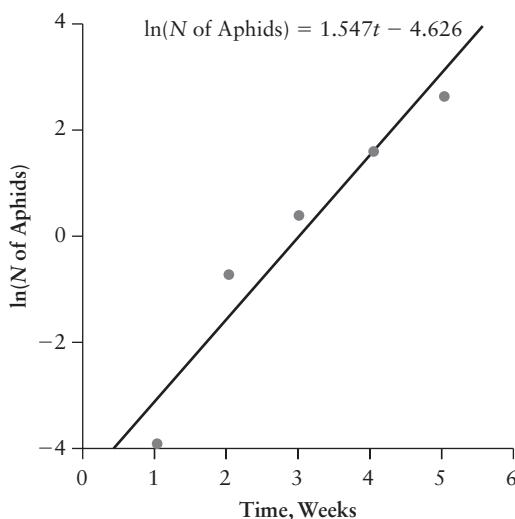


FIGURE 1.2. Plot of aphid data (from table 1.1).

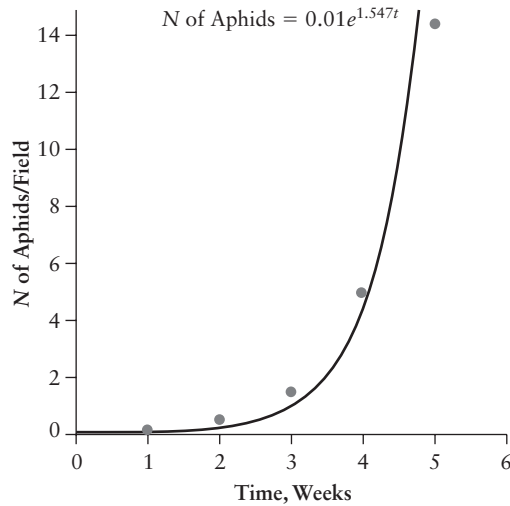


FIGURE 1.3. Aphid data from figure 1.2 plotted arithmetically.

which can be rearranged as

$$t = [\ln(N \text{ of aphids}) + 4.626] / 1.547.$$

The natural log of 40 is 3.69, so we have

$$t = (3.69 + 4.626) / 1.547 = 5.375.$$

Translating this number into the actual date (remember that April 22 was time = 5), we see that the critical number will arrive about April 24 (at 3:00 p.m. on April 24, theoretically).

In practice, calculations like these exclude many complicating factors, and we should never take too seriously exact predictions the model makes. On the other hand, April 24 really does represent the best prediction we have based on available data. It may not be a very good prediction, but it is the only one available. Furthermore, it may seem quite counterintuitive that, having taken five full weeks to arrive at only 14 aphids per plant, in only two more days the critical figure of 40 aphids per plant will be realized—such is the nature of exponential processes. A simple model like this could help the farmer plan pest control strategies.

EXERCISES

- 1.7 Plot the derivative (which you can approximate by subtracting the current N from its successive one and dividing by the change in time) as a function of N for the data calculated in exercise 1.6 (use a very small time interval, e.g., 0.01).
 - 1.8 From the data calculated in exercise 1.7, divide the estimated derivative by the average number of individuals during the time period, and plot them over time. What is the result, and what does it mean?
-

Density Dependence

In the previous section we showed that any population reproducing at a constant per capita rate will grow according to the exponential law. Indeed, that is the very essence of the exponential law: each individual reproduces at a constant rate. However, the air we breathe and the water we drink are not completely packed with bacteria or fungi or insects, as they would be if populations grew exponentially forever. Something else must happen. That something else is sometimes referred to as intraspecific competition, which means that the performance of the individuals in the population depends on how many individuals are in it. More generally, it is referred to as density dependence. It is a complicated issue that has inspired much debate and acrimony in the past and still forms an important base for more modern developments in ecology.

The idea was originally associated with the human population, and was brought to public attention as early as the nineteenth century by Sir Thomas Malthus (1960 [1830]). It was formulated mathematically by Verhulst (1838) as the “true law of population” (Doubleday 1842); Verhulst’s version is better known today as the logistic equation (see below). Later, Pearl and Reed (1920), in attempting to project the human population size of the United States, independently derived the same equation. Associated with its mathematical formulation was a series of laboratory studies with microorganisms in the first part of the twentieth century. Most notable were those of Gause (1934), in which population growth was studied from the point of view of competition, both intra- and interspecific.

In the early part of the twentieth century a variety of terms were introduced, all of which essentially referred to the same phenomenon of approaching some sort of carrying capacity through a differential response of per capita population growth rate to different densities. Chapman (1928) formulated the idea in terms of “environmental resistance.” Howard and Fiske (1911) categorized mortality factors as either “catastrophic,” in which some proportion of the population died regardless of its density, or “facultative,” which caused an increasing proportion of the population to die with the increasing density of the population. In 1928 Thompson redefined *catastrophic* as “general” and “independent” of density and *facultative* as “individualized” and “dependent” on density, and later Smith (1935) proposed the density-dependent/density-independent gradient. Thus by the 1930s the dichotomy of density independence versus density dependence had taken firm root after having been sown not long after the turn of the century.

In the 1930s Nicholson (1933) and his colleague Bailey (Nicholson and Bailey 1935) formalized the concept of regulation through density-dependent factors and clearly associated the idea of intraspecific competition with density dependence. In Nicholson and Bailey’s conceptualization of density dependence, four points were proposed: (1) population regulation must be density dependent; (2) predators and parasites, as well as intraspecific competition, may function as density-dependent forces; (3) more than density dependence alone may function to determine actual population size; and (4) density dependence does not always function to regulate population density.

In contrast to the ideas of Nicholson and Bailey (and especially of their later followers) were the ideas of Andrewartha and Birch (1954). They held that the environment is not divisible into density-dependent versus density-independent forces and that, although resources might limit populations, they rarely do so because some aspect of the physical environment (usually collectively referred to as “the weather”) almost always imposes its effects. They furthermore noted that the mathematical models that presume equilibrium and persistence are not really necessary if there is indeed no “balance” in nature (density dependence strongly implying some sort of balance of nature). Most data sets failed to support the idea of density dependence, and the idea was viewed as possibly untestable. Rather, Andrewartha and Birch argued, the regulation of populations was frequently taken as an article of faith. The problem was this: how long could a population persist without regulation? The researchers’ recognition of the fact that local populations would frequently go extinct but would be refounded from other population centers anticipated ideas of metapopulations that would become popular some 20 years later.

Milne (1961) modified both versions of population regulation and noted that perfect density dependence, if it ever exists, does so only at very high densities. Rather, what most characterizes populations in nature is what might be referred to as imperfect density dependence, whereby predators and parasites plus density-independent effects usually hold populations below levels at which intraspecific competition can become important. More generally, Dempster (1983) suggested that density independence can be operative within limits, such that an upper ceiling will be imposed on the population and a lower limit will prevent the population from going extinct. Strong (1986) referred to such regulation within bounds rather than around a particular deterministic value as “density vagueness.” These variations are fundamentally in the density-dependence camp but with strong notions of nonlinearities and the importance of spatial distribution, topics discussed in later chapters.

In the end it would seem that the entire debate about density-dependent versus density-independent control of populations was focused on a false dichotomy. In a variety of guises (e.g., intermediate disturbance, metapopulations) modern ecology has come to acknowledge that both density-dependent and density-independent forces may function together to regulate populations in nature. But, more important, there is general agreement that the dynamic behavior of a population (its change in numbers over time) does not necessarily suggest any particular mechanism of regulation. Over the past two decades a large literature has developed that seeks to use advanced methods of analysis to determine whether density dependence operates (see, for example, Hastings et al. 1993), but that is beyond the intended scope of this text. Part of this later literature is associated with the possibility that many populations under density-dependent control actually may be chaotic (discussed more fully in chapter 4). Chaotic populations can easily be confused with random populations, and it has been noted that one way of resolving some of the earlier debates about the issue is to acknowledge that extreme density dependence

(which would promote chaos) could easily produce population behavior that looked quite density independent (i.e., chaotic) (Gukenheimer et al. 1977).

As one can see from the previous paragraphs, the literature on density dependence is enormous. Yet much of it can be divided conceptually into three categories. First, the effect of density on the growth rate of the population (be it through declining reproduction or increased mortality) is simply added to the exponential equation to form the famous logistic equation (as discussed below). Traditionally the logistic equation is expressed in continuous time as a differential equation, but recently there has been considerable interest in the special properties of the logistic idea expressed in discrete time, the logistic map. The logistic equation, in either its continuous or its discrete form, treats the population growth rate as a single constant, even though we understand that it actually represents birth rate minus death rate. Other approaches treat each of these rates separately.

Decomposing the population growth rate into its two components, the second category of literature focuses on the relationship between density and reproduction (i.e., on the fact that density modifies birth rate). We guess that the first acknowledgment of density dependence in nature was probably by the world's first farmers. When planting crops they soon realized that higher planting densities provide higher yields (which, in principle, are correlated with reproductive output), but only to a point. Once a high enough density is reached, further increases in density fail to provide further increases in yields. This general relationship is referred to as the yield–density relationship and is, in some respects, the most elementary form of density dependence. Originally developed mainly in the agronomy literature, the relationship between density and yield subsequently became an important theoretical baseline for general plant ecology. Yield was usually seen as a product of reproduction of plants such as corn and soybeans, and thus the subject of yield and density can be thought of more generally as the relationship between density and reproduction.

Finally, the third topic is the possibility that density affects survivorship rather than reproduction (i.e., that density modifies death rate). The main literature on this topic was originally conceived in the area of forestry, but it has since become generalized under the subject of self-thinning laws, which is used mainly in plant ecology but also in fisheries research. The yield–density relationship, discussed in the previous paragraph, involves examining yields of different populations that have been sown at different densities. It is a static approach in this sense. Once established through sowing, the population density remains constant and the variable of interest is the yield. An alternative approach is more dynamic and follows changes in both size (biomass) and density over time in the same population. This more dynamic approach considers mortality as well as growth, and in the context of forestry, where it was originally developed, mortality is known as thinning.

In the following three sections we follow this basic schema: (1) the logistic equation, (2) yield–density relations, and (3) self-thinning laws. But first we present some exercises to set the stage for the introduction of density dependence.

EXERCISES

- 1.9** The following are UN data on the global human population (in hundreds of thousands) from 1964 through 2007 (United Nations, Department of Economic and Social Affairs, 2011):

Year	Density	Year	Density
1964	3,278	1985	4,844
1965	3,347	1986	4,927
1966	3,417	1987	5,013
1967	3,486	1988	5,099
1968	3,558	1989	5,185
1969	3,632	1990	5,273
1970	3,708	1991	5,357
1971	3,785	1992	5,440
1972	3,861	1993	5,521
1973	3,936	1994	5,601
1974	4,011	1995	5,681
1975	4,084	1996	5,762
1976	4,156	1997	5,840
1977	4,226	1998	5,918
1978	4,298	1999	5,995
1979	4,372	2000	6,072
1980	4,447	2001	6,147
1981	4,522	2002	6,222
1982	4,601	2003	6,297
1983	4,682	2004	6,373
1984	4,762	2005	6,449

Estimate the intrinsic rate of increase (or decrease) for each successive year, and plot the estimated intrinsic rate versus the population density. Is it a flat, straight line, as would be expected if the population were growing exponentially? What kind of function would describe the data relatively well?

- 1.10** Graphically estimate the point at which the human population will stop growing (the number of individuals in the world at the point at which the present trend will reach zero. How would various sections of the data, if considered alone, change your conclusions? What if you had made your estimate in 1976 based on the most recent 10 years of data?
- 1.11** Using the UN data on the human population, insert the linear approximation of the intrinsic rate of natural increase as it relates to the population density, and graphically estimate the value of the intrinsic rate of increase (the limiting value as the density approaches zero).

- 1.12** Beginning with the exponential equation, assume that the rate of increase declines linearly with population density (i.e., substitute $a - bN$ for r). Show what substitutions you need to make to write the equation in its more biologically sensible form (i.e., with the intrinsic rate of natural increase and the carrying capacity evident). How do a and b of the linear approximation relate to the r and K of $\frac{dN}{dt} = rN\frac{K - N}{K}$?
-

The Logistic Equation

Density dependence is generally regarded as the major modifier of the exponential process in most populations. Consider, for example, the data shown in figure 1.4 (Vandermeer 1969). The protozoan *Paramecium bursaria* was grown in bacterial culture in a test tube, and the data shown are for the first 15 days of culture (the data are the number of cells per 0.5 ml). In figure 1.5 those numbers are shown as a graph of $\ln N$ versus time (recall how the intrinsic rate of natural increase was estimated in this way).

The relationship is approximately linear (see figure 1.5), and our conclusion would be that the population is growing according to an exponential law. If this equation were followed into the future, we would have a very large population of *Paramecium*. Indeed, considering the size of *P. bursaria*, about 3,000 individuals could fit into 0.5 ml if you stacked them as sardines. Thus the 3001st individual would cause all the animals to be squeezed to death, recalling the leitmotif with which this chapter began. We can compute exactly when this would happen from the equation of the line in figure 1.5, which is

$$\ln(N) = 1.239 + 0.227t.$$

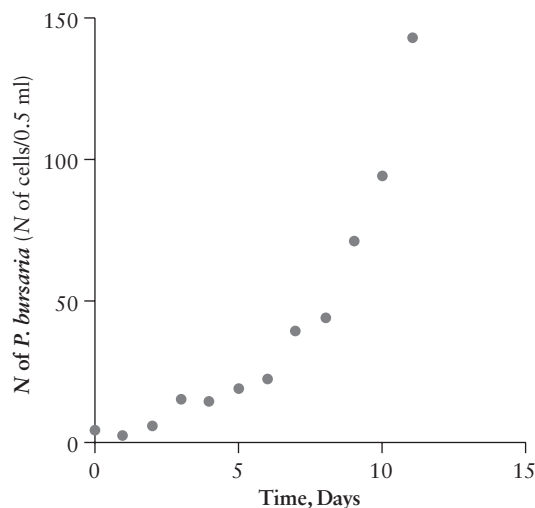


FIGURE 1.4. Growth of a culture of *Paramecium bursaria* in a test tube (Vandermeer 1969).

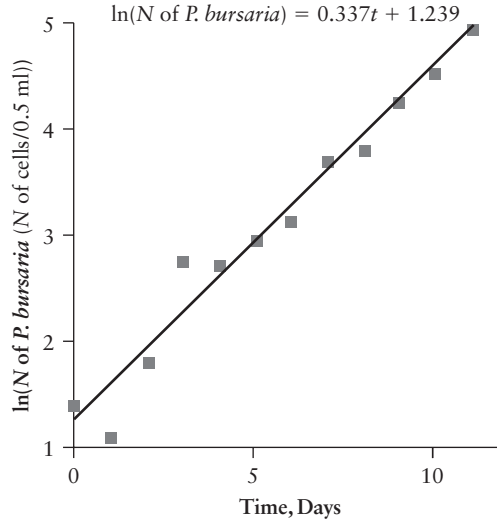


FIGURE 1.5. Logarithmic plot of the data from figure 1.4.

We substitute the critical value of 3,001 individuals to obtain

$$\ln(3,001) = 0.227 + .337t,$$

which can be arranged to read

$$t = \frac{[\ln(3,001) - 1.239]}{0.337} = 20.80.$$

Thus, on the basis of an 11-day experiment we can conclude that after about 21 days, the test tube will be jam-packed with *P. bursaria*, such that all the individuals will suddenly die when that 3,001st individual is produced. The actual data for the experiment carried out beyond the 24-day expected protozoan Armageddon are shown in figure 1.6.

These data suggest that something else happened. As the density of the *Paramecium* increased, the rate of increase declined, and eventually the number of *Paramecium* reached a relatively constant number. The theory of exponential growth must be modified to correspond to such real-world data.

Let us begin with the exponential equation but assume that the intrinsic rate of growth is directly proportional to how much resource is available in the environment. Thus we have

$$\frac{dN}{dt} = rN, \tag{12}$$

the classical exponential equation discussed earlier in this chapter. But here we presume that r is directly proportional to F (i.e., $r = bF$), where F is the amount of resources (F for food) in the system that is available to the population. Thus, equation 12 becomes

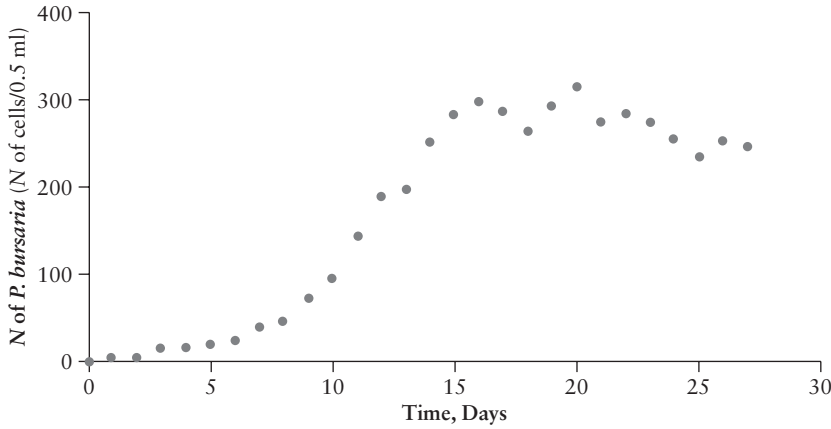


FIGURE 1.6. Long-term data for the same populations of *Paramecium bursaria* (data from time = 0 to time = 12 are the same as in figure 1.4).

$$\frac{dN}{dt} = bFN. \quad (13)$$

But now we assume that there is no inflow of resources into the system so that the total amount of resources is constant and is divided up into the part that is usable by the population and the part that has already been used. That is,

$$F_T = F + cN, \quad (14)$$

where F_T is the total amount of resources in the system and c is the amount of F held within each individual in the population. Equation 14 can be manipulated to read

$$F = F_T - cN. \quad (15)$$

Substituting equation 15 into equation 13, we have

$$\frac{dN}{dt} = b(F_T - cN)N, \quad (16)$$

whence we see that equation 16 is a quadratic equation. Finding the equilibrium points—that is, the points at which the population neither increases or decreases—is done by setting the derivative equal to zero, thus obtaining

$$0 = b(F_T - cN)N,$$

which has two solutions. The first solution is at $N = 0$, which simply says that the rate of change of the population is zero when there are no individuals in the population. The second solution is at F_T/c , which is the maximum value that N can have. This is the value of N for which $F = 0$ (when all the resources in the system are contained within the bodies of the individuals in the population). Because the limitations of the environment are more or less stipulated by the value of F_T and the maximum number of individuals that

the environment can contain is F_T/c , the value F_T/c is frequently referred to as the carrying capacity of the environment (the capacity of the environment to carry individuals). The traditional symbol to use for carrying capacity is K , so we write $K = F_T/c$. We also note that as the population approaches zero (as N becomes very small but not exactly at zero), the rate of increase of the original exponential equation will be bF_T (because the general equation is bF , and when N is near zero F is almost the same as F_T). After some manipulation of equation 16 we can write

$$\frac{dN}{dt} = bF_T N \left(\frac{\frac{F_T}{c} - N}{\frac{F_T}{c}} \right).$$

Now, substituting r for bF_T and K for F_T/c , we obtain

$$\frac{dN}{dt} = rN \frac{K - N}{K}, \quad (17)$$

which is the classic form of the “logistic equation.” Note the form of the equation. It has a very simple biological interpretation. The quantity $(K - N)/K$ is the fraction of the carrying capacity that has not yet been taken up by the individuals in the population. In shorthand we might refer to this quantity—the fraction of the carrying capacity or the fraction of total available resources—as the available resource space. Then the logistic equation is obtained by multiplying the original intrinsic rate of increase, r , by the available resource.

Returning to the earlier example of *Paramecium bursaria*, a glance at the data suggests that the carrying capacity is around 290 individuals (averaging all the points after the data have leveled off). The original estimate of r as 0.337 was probably too low (because the effects of density dependence were probably effective even during the time of the initial growth), so taking a slightly larger value, let $r = 0.5$. The logistic equation for these data then becomes

$$\frac{dN}{dt} = 0.5N \frac{290 - N}{290},$$

which is plotted in figure 1.7, along with the original data. This example represents a reasonably good fit to the logistic equation.

The existence of density dependence also calls into question the extrapolations that one is tempted to make from a process that seems inexorably exponential. The example earlier in this chapter of the aphids in the cornfield is a case in point. Concluding that the farmer had only two days before disaster struck may have been correct, but it also could have been grossly in error, depending on the strength of the density dependence. Indeed, with strong density dependence, the field’s carrying capacity for the herbivore could have been below the threshold at which the farmer needed to take action, in which case no action at all would have been necessary.

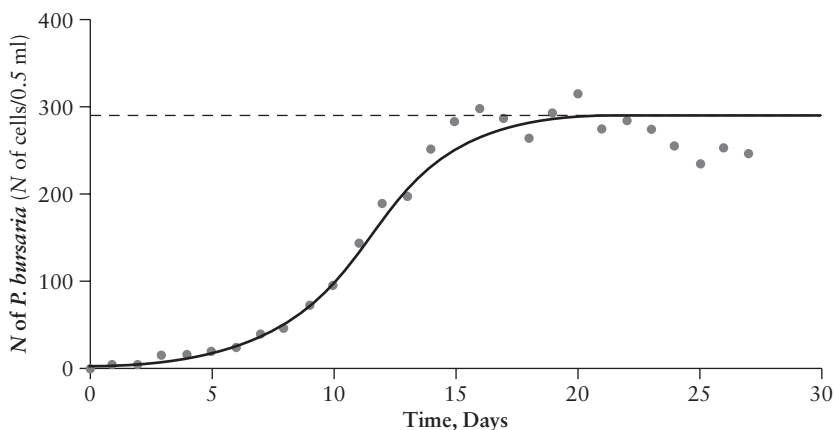


FIGURE 1.7. Fit of logistic equation to the *Paramecium bursaria* data from figure 1.6.

In some management applications, (e.g., fisheries) it is desirable to maximize the production of a population, which is to say, to maximize the rate of increase, not the actual population. The logistic equation can provide a useful guideline for such a goal because it is reasonably simple to calculate what population density will produce the maximum rate of population increase (see exercise 1.13). Thus, once the carrying capacity is known, the population density at which the rate of growth will be maximized is automatically known. In actual practice this so-called maximum sustained yield has some severe problems associated with it, largely stemming from the simplifying assumptions that go into its formulation (these issues are more fully discussed in chapter 4).

EXERCISE

- 1.13** Assuming that a population grows according to the logistic equation, what is the value of the population density that will give the maximum growth rate?
-

The Yield-Density Relationship

The process of intraspecific competition or, more generally, density dependence is certainly extremely common, if not ubiquitous, and thus legitimately calls for a theoretical framework, the most common and general of which is the logistic equation. However, for many applications it is not sufficient to consider only the population growth rate but rather is necessary to decompose that rate into its component parts, birth rate and death rate. In this section we consider the effect of density on birth rate. This theory was developed mainly from work on plants, especially in agroecosystems. Farmers need to know the relationship between planting density and the yield of a crop (which is frequently the seed output). This relationship is known as the

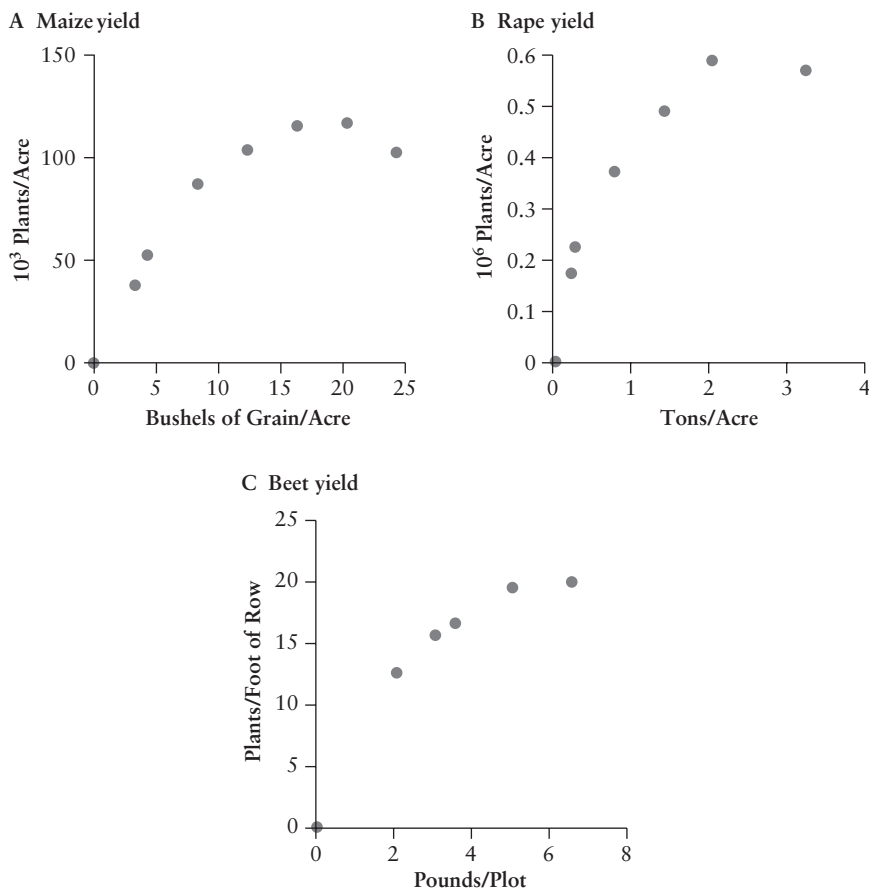


FIGURE 1.8. Exemplary yield versus density data for maize, rape, and beets (from Willey and Heath 1969).

yield–density relationship and is the basis of much agronomic planning as well as a springboard for much general plant ecology. For our purposes here, the yield–density relationship provides the most elementary form of the effect of intraspecific competition on reproduction and lays bare its essential elements. We thus give considerable space to the development of the principles of intraspecific competition as reflected in the yield–density relationship.

The formal elaboration of yield–density relationships probably first appeared in 1956 with the work of Shinozaki and Kira. These researchers noted, as had many before them, that plotting yield versus density for various plant species usually results in a characteristic form. Several examples are shown in figure 1.8.

Shinozaki and Kira suggested a simple hyperbolic form,

$$\frac{Nw_{\max}}{1 + aN},$$

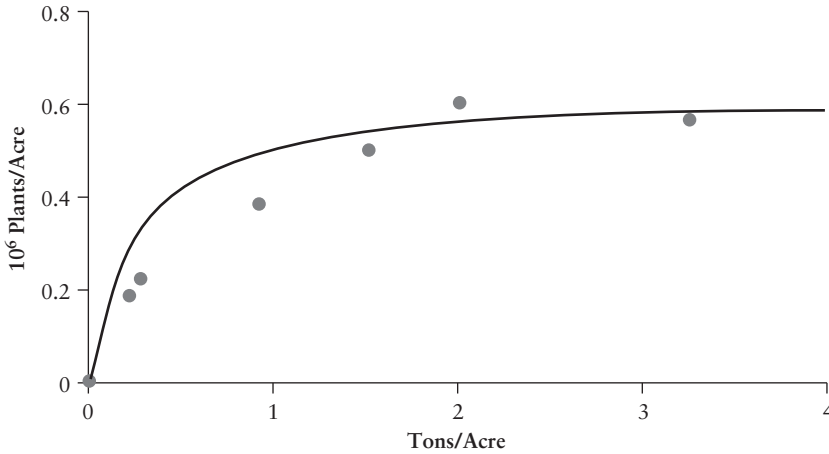


FIGURE 1.9. Equation of Shinozaki and Kira superimposed on the rape data from figure 1.8.

where N is population density, Y is yield, $w_{\max}$ is the unencumbered (i.e., without competitive effects) yield of an individual plant, and a is an arbitrary constant. This equation asymptotes as N becomes very large and thus corresponds to another well-known empirical observation in plant ecology known as the law of constant final yield (which actually is not always true, as discussed below). In figure 1.9 Shinozaki and Kira's equation is shown in relation to the data for rape in figure 1.8.

This empirical curve fitting can be rationalized with some simple plant competition theory. We begin by considering what might happen with individual plants and later accumulate those plants into a population so as to examine the effect of density. Consider a single corn plant in a pot. When provided with all the necessary light, water, and nutrients, it will grow to some specified height with some specified biomass. If two corn plants are planted in a pot of the same size and provided with the same amount of light, water, and nutrients, each of the plants will attain a biomass smaller than that of the corn plant grown alone because the same amount of resources is being used by two individuals rather than one. If we symbolize the biomass a plant attains when growing alone as k , we can write the simple linear relationship

$$w_1 = k - \alpha w_2, \quad (18)$$

where w_1 refers to the biomass (weight) of the first plant, w_2 refers to the biomass of the second plant, and α is the proportionality constant that expresses the decrease in biomass of the first individual as a proportion of the biomass of the second individual. This makes the very reasonable assumption that larger plants have greater competitive effects so that competitive effects are directly proportional to their size. Rearranging equation 18, we see that

$$\alpha = (k - w_1)/w_2. \quad (19)$$

This same development could be applied to three plants growing in a single pot, in which case the equation describing the results would be

$$w_1 = k - \alpha_{12}w_2 - \alpha_{13}w_3, \quad (20)$$

where α_{12} is the effect of a unit of biomass of individual 2 on the biomass of individual 1 and α_{13} is the effect of a unit of biomass of individual 3 on the biomass of individual 1. The parameter α is frequently referred to as a competition coefficient because it represents the effect of one individual on another. The calculation of α from real data is quite easy when we have only two plants: grow a single plant in a pot, and measure its biomass after some specified time, giving the value of k ; then grow two plants in a pot, and measure their biomasses, giving the values of w_1 and w_2 ; then apply equation 18 to determine the value of α . The estimation of the competition coefficients when there are more than two individuals is somewhat more complicated but need not concern us at this point. For now it is important only to understand the logic of the thinking that went into the construction of equations 19 and 20. We now proceed to generalize equation 19.

Let us suppose that instead of planting just two or three individuals in a pot, we plant a large number of individuals. If the total number planted is n , we can expand equation 20 by simply adding more terms until we have added all n individuals to the calculation. That is, equation 20 for n individuals becomes

$$w_1 = k - \alpha_{12}w_2 - \alpha_{13}w_3 - \dots - \alpha_{1n}w_n$$

or, more compactly,

$$w_1 = k - \sum \alpha_{1j}w_j, \quad (21)$$

where the summation is taken from $j = 2$ to $j = n$. If all the individuals are exactly the same, it might be argued that all the α_{1j} values are equal. As a first approximation, this is probably a good assumption. However, there is a crucial way in which the competition coefficients differ from one another, as becomes evident below when we try to elaborate this same example from the level of a pot to the level of a field.

For now, assume (a bit unrealistically) that all individuals produce the same biomass and that the competition between any two pairs of individuals is identical from pair to pair (i.e., assume we can substitute the mean values for biomass and the competition coefficients). We can thus write

$$\bar{w} = k - N\bar{\alpha}\bar{w}, \quad (22)$$

where N is the population density, $\bar{\alpha}$ is the mean value of the competition coefficients, and $\bar{w}$ is the mean biomass. Equation 22 can be rearranged as follows:

$$\bar{w}(1 + N\bar{\alpha}) = k,$$

and finally

$$\bar{w} = \frac{k}{(1 + N\alpha)}, \quad (23)$$

If $\bar{w}$ is the biomass of an average individual in the population, the total population yield must be

$$Y = \bar{w}N.$$

Substituting from equation 23 for $\bar{w}$, we obtain

$$Y = \frac{Nk}{(1 + \alpha N)}, \quad (24)$$

which is identical to the empirical equation of Shinozaki and Kira. The advantage of equation 23 is that because of the derivation based on plant competition theory, the parameters in the equation have obvious meaning; k is the unencumbered yield of an average individual plant, and α is the mean competition coefficient between two individual plants.

An additional complication arises when we have data like the maize data in figure 1.8, where at high densities the yield actually falls. To accommodate data such as these Bleasdale and Nelder (1960) suggested modifying the basic Shinozaki and Kira equation with an exponent, citing either

$$Y = \frac{Nk}{(1 + \alpha N)^b} \quad (25)$$

or

$$Y = \frac{Nk}{(1 + \alpha N^b)} \quad (26)$$

as a reasonable approximation of data that are shaped parabolically. The constant b is, in the context of Bleasdale and Nelder's derivation, a fitted constant that they presume is related to an allometric effect (i.e., the harvested material is produced proportionally less at higher densities of the plant). Either equation reduces to Shinozaki and Kira's equation when $b = 1.0$. Bleasdale and Nelder chose the first of these two equations arbitrarily, and it has become something of a standard in plant ecology literature. It is worth noting that it is not only the allometric effect that can produce a yield–density curve that descends at high densities (see exercise 1.14).

The analyses in this section have been based on static descriptions of the consequences of different initial densities or reproductive yield within a single season. However, because reproductive yield in one year can be expressed as the number of offspring and therefore as the population size in the next time period (rather than biomass, as assumed so far), this formulation can be easily converted to a model of density-dependent population dynamics (Watkinson 1980), as shown in exercise 1.15 and discussed at the end of this chapter.

EXERCISES

- 1.14** The derivation of the Shinozaki and Kira equation presumed that competition remains constant among individuals. Yet basic plant ecology might suggest that as individual plants become closer to one another, their competitive effects will become more severe. Assume a declining relationship between competition and distance between individual plants, translate that assumption into a relationship between α and N , and substitute into the Shinozaki and Kira equation. What results?
- 1.15** In the Bleasdale and Nelder equation, the “yield” of the equation effectively represents the input into the next generation if, instead of total crop yield in terms of biomass, it is expressed as the number of offspring. Write an equation, using the basic Bleasdale and Nelder formulation, that relates the population density this year to that we expect next year, assuming 100% mortality (i.e., an annual species). If you plot the number this year on the x axis and the number next year on the y axis, what does the functional form look like for various values of the parameter b ?
-

Density Dependence and Mortality: Thinning Laws

In the above developments we assumed that density dependence acts in such a way that the growth of individuals is slowed by a larger population and that a decline in individual growth rate reflects a lower birth rate that eventually stabilizes the population at some particular number. In our explanation of the logistic equation, we made no explicit assumption about birth and/or mortality, and the derivation revolved around the intrinsic rate of natural increase, which includes both death and birth rates. However, implicitly in the section on the logistic equation and explicitly in the above section on the yield density–relationship, the assumption was that we were dealing exclusively or mainly with birth rate modifications rather than death rate modifications. There are times where the distinction can be crucially important. For example, the growth in total biomass of a plantation of trees is usually approximately logistic in form, but the same logistic equation could describe the pattern in either figure 1.10A or 1.10B. The difference between the two figures is not trivial from a forester’s point of view. In figure 1.10A there are large numbers of very small trees, none of which is harvestable, while in figure 1.10B there is a smaller number of larger trees. The point is that in figure 1.10A there has been a great deal of intraspecific competition, but it took the form of each individual growing more slowly and almost no mortality, while in figure 1.10B one of the main responses to intraspecific competition was for some individuals to die while others continued growing rapidly. The biomass of the forests in both figures is the same (that is the way the example was constructed), but one forest will be useful for harvest and the other not. Similar examples could be given for any organism with indeterminate growth. For example, many fish become stunted when in very dense populations and thus represent less of an attraction for sport fishers and can have significant consequences for commercial fisheries as well.

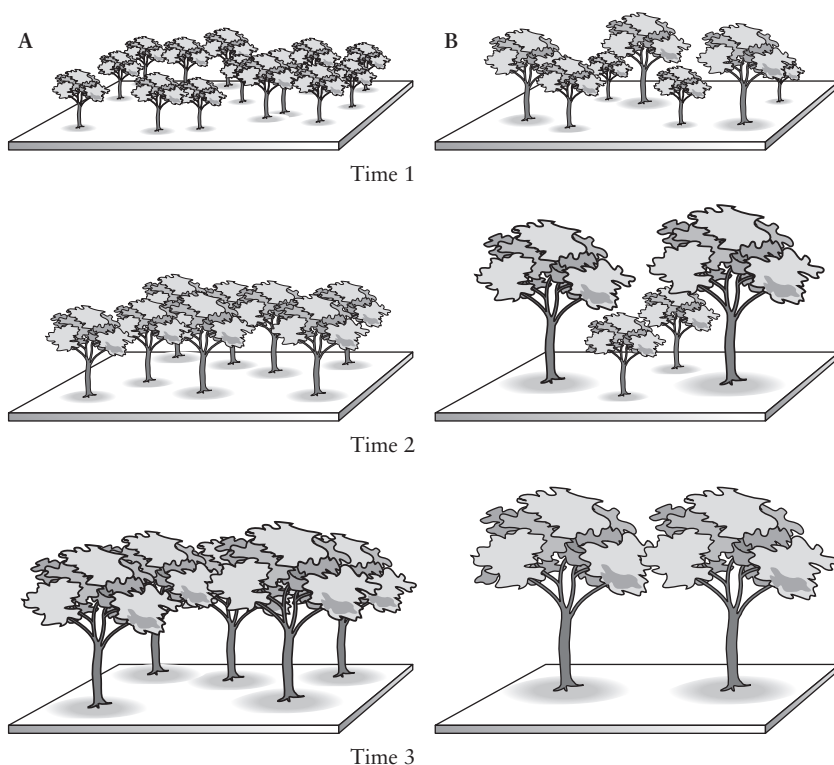
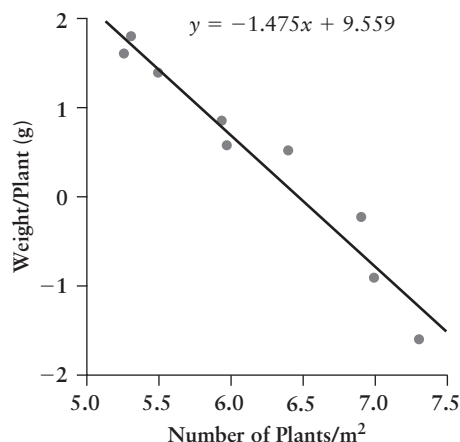


FIGURE 1.10. Diagrammatic representation of the process of thinning (see text).

Reflect for a moment on the pattern of growth and mortality in a densely planted tree plantation or in a natural forest when large numbers of seeds germinate more or less simultaneously. First, seedlings are established at a very high density. Walking through a beech-maple forest, for example, one is struck by the carpet of maple seedlings in almost every light gap one encounters. As the seedlings grow, the increase in biomass of each individual treelet is limited by intraspecific competition, as described above. But when populations are sufficiently dense, inevitably some individuals come to “dominate” (be larger) while others become “suppressed” (remain small due to competition from their neighbors). Eventually the suppressed individuals die, and we say that the population has been thinned. But then the trees keep on growing and the process repeats itself, with some suppressed and others dominating. In this way a population of plants that began at a very high density is thinned to the point that the adults are at some sort of carrying capacity. In some ways this process seems to be the reverse of what was described above in the development of the logistic equation. Here we begin with a number larger than K , and through the process of thinning the population is reduced to K , rather than beginning with a small population and increasing to the value of K . On the other hand, remember that biomass is increasing throughout the process.

FIGURE 1.11. Typical relationship between log density of surviving plants and log dry weight. Example is of *Helianthus annuus* (Hiroi and Monsi 1966).



This phenomenon is most easily seen as a graph of the log of individual biomass versus the log of density at successive harvests of the same population over time, as shown with the data in figure 1.11 and more schematically in figure 1.12. Although plotting individual mass rather than total population yield, this seems similar to the analysis of density effects on birth rate in the previous section. But figure 1.11 is much more dynamic; we start with a single starting density and observe changes in biomass (or some related variable). If no mortality occurs, we expect a straight vertical line, that is, the per-plant biomass increases but the population density remains constant (see figure 1.12A). But if there is mortality, over time the curve will shift to the left, to lower densities, while at the same time the per-plant biomass will increase (see figure 1.12B). If, on the other hand we had begun with two different populations at slightly different densities we would see that both populations would increase in biomass. Further, assuming that densities were such that this initial increase in biomass happened without competition, both populations will grow in biomass by the same amount (see figure 1.12C). If we let both of these populations continue to grow, we can expect some thinning (mortality) to occur, especially in the more dense population (see figure 1.12D).

Here we can see both a plastic effect on growth and a mortality effect. The plastic effect is seen as a smaller biomass increase at larger densities, as shown in figure 1.12D. The mortality effect is seen as a decrease in density at higher densities, as shown in figure 1.12D. If we continue the pattern of development illustrated in figure 1.12 through time, we see that each population begins its process of thinning as it approaches a theoretical thinning line, as can be seen in the data shown in figure 1.13. Once mortality starts, the population tends to follow a straight line on a log-log scale. This kind of relationship has been shown many times—most often in herbaceous plants over time or in woody plants, comparing plantations at different densities.

This self-thinning law (*self-* because no forester or agronomist is there doing it) was first developed in plants, but many animals show a similar pattern. It is a very nice way of showing the growth and mortality effects of

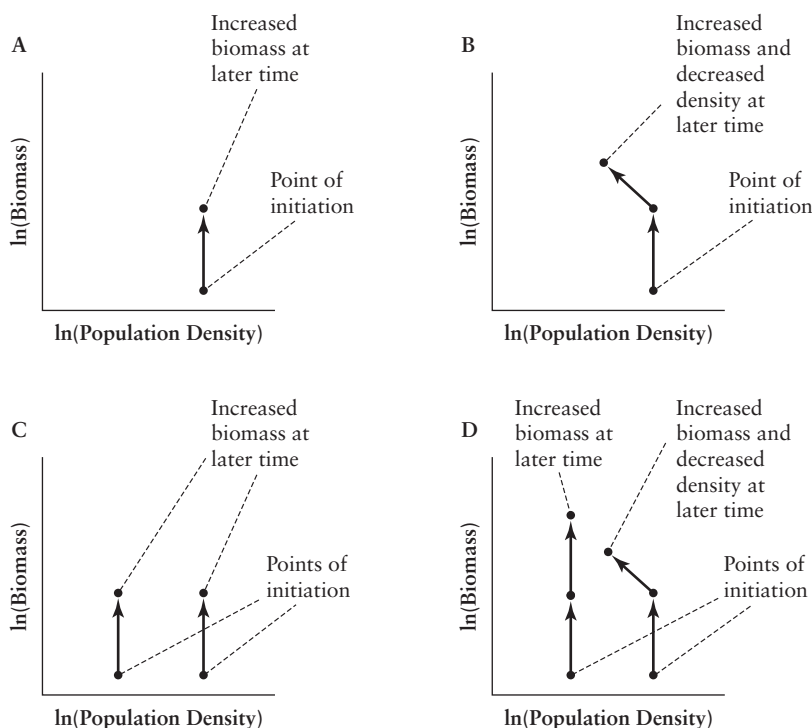


FIGURE 1.12. Expected pattern of growth and mortality as the thinning process and growth in biomass interact. $\ln(\text{biomass})$ refers to the natural logarithm of the biomass of an average individual in the population.

density in the same graph. But it also provides an elegant way of looking at density-dependent mortality that can be easily compared among species on very different time scales because time is not explicit. Furthermore, some time ago plant ecologists noticed that this process of self-thinning always seems to take on a particular pattern. When plotting the logarithm of the biomass of an average individual plant versus the population density at the time the biomass is measured, the points in a thinning population appeared linear; furthermore, the slope of the line always appeared to be nearly $-3/2$ (e.g., see figures 1.12 and 1.13), provided that the population was undergoing thinning. This phenomenon has come to be known as the three-halves thinning law.

Yoda and colleagues (1963) provided an elegant theory explaining the origin of the law. Suppose that each plant is a cube. If each side of the cube is x , the area of one of the cube's faces is x^2 and the volume of the cube is x^3 . Now imagine that the plantation is made up of a large number of these cubes and that they begin growing and thinning through intraspecific competition. The overall process is illustrated in figure 1.14. The area of the plantation is A . The population density is the total area divided by the surface area occupied by a single plant (that is, a single cube). Thus N , the population density, is equal to A/x^2 . Now presume that the biomass, w , of an individual plant is

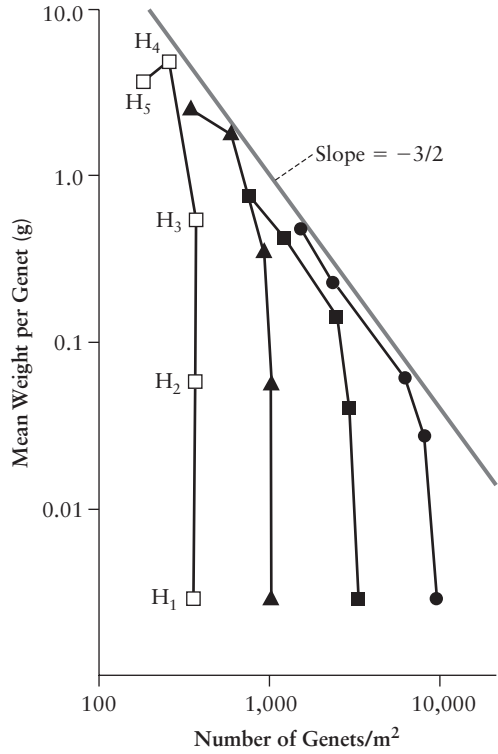


FIGURE 1.13. Relationship between the density of genets and the mean weight per genet in populations of *Lolium perenne*. H_1 , H_2 , and so on are successive harvests (Kays and Harper 1974).

approximately equal to the volume of the cube representing it, so that $w = x^3$. So we have the pair of equations

$$N = A/x^2$$

and

$$w = x^3.$$

Rearranging these equations, we write

$$x = A^{1/2}N^{-1/2}$$

and

$$x = w^{1/3}.$$

Because the left-hand side of both equations is equal to x , we can set the right-hand sides of both equal to each other, giving

$$A^{1/2}N^{-1/2} = w^{1/3},$$

which simplifies to

$$w = A^{3/2}N^{-3/2},$$

which can be put in the more standard form

$$\ln(w) = (3/2)\ln(A) - (3/2)\ln(N),$$

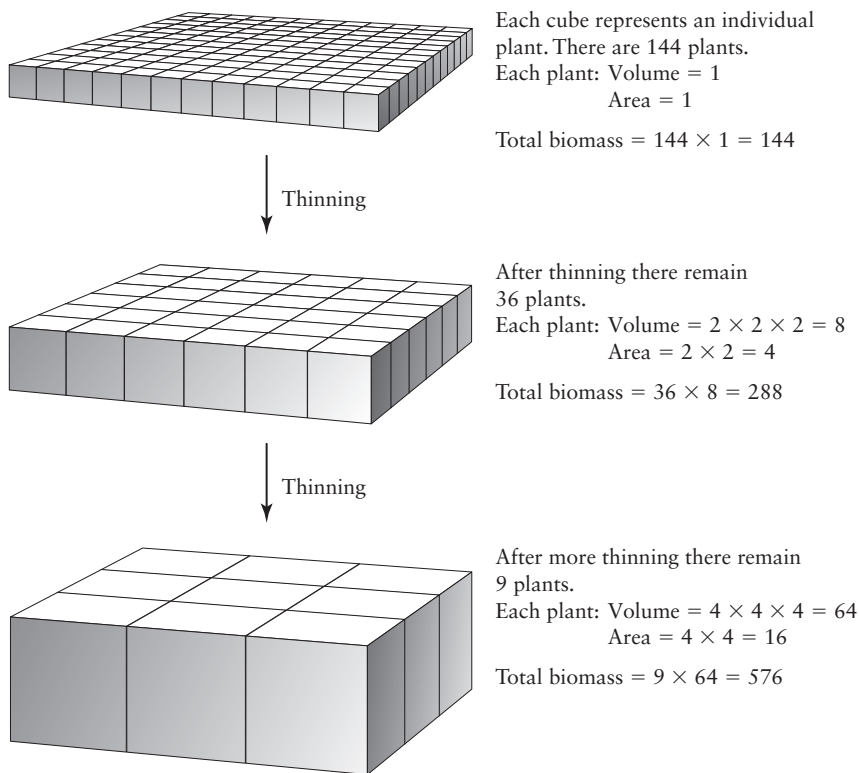


FIGURE 1.14. Yoda and colleagues' (1963) interpretation of the origin of the three-halves thinning law.

which represents a straight line with a slope of $-3/2$ on a graph of $\ln(w)$ versus $\ln(N)$.

Thus we see from simple geometric reasoning that it is not unusual to expect the three-halves thinning law. On the other hand, the basic empirical base of the “law” has been persistently questioned (e.g., Westoby 1984). In fact, when large numbers of data sets are examined, the slope has been shown to be closer to $-4/3$. And a prediction of $-3/2$ from geometric considerations is actually quite shaky—plants are not, after all, cubes.

Current thinking has largely moved away from purely geometric considerations. A quite different explanation and quantitative prediction of thinning relationships was provided by Enquist et al. (1998) based on fundamental metabolic scaling relationships for both plants and animals. The metabolic model is based on the observation that the resource use and metabolic rate of both plants and animals tend to increase with body size with a power of $-3/4$ (West et al. 1997). Assuming that plants grow until they are resource limited and that resources limit the total productivity (biomass $\times$ density) of an area, this model predicts a thinning slope of $-4/3$, which is consistent with more recent and extensive data (Enquist et al. 1998).

Density Dependence in Discrete Time Models

Using much of the same qualitative reasoning as above, the process of density dependence can be formulated in discrete time rather than continuous time. Rather than asking how a population grows instantaneously, we can ask how many individuals will be in the population next year (or in other time unit) as a function of how many are here now. Recall equation 1.3, the exponential equation,

$$N_{t+1} = \lambda N_t,$$

which is a statement of population growth in discrete time. Now, rather than proceeding with a generalization about future numbers (which was the direction taken earlier), we remain in the realm of discrete time and ask what modifications might be necessary to make this equation density dependent. In other words, what do we come up with if we use the same rationale we used in developing the logistic equation but this time do it in discrete time?

It seems reasonable to suppose that the population will grow slowly if it is near its carrying capacity (K) and will grow more rapidly if it is far below its carrying capacity. This is the same as saying that λ varies with population density. If we simply allow λ to vary linearly with density (i.e., let $\lambda = r - bN_t$, precisely the same *conceptual* approach we took with the logistic equation), we write

$$\lambda = r - bN_t = \lambda \frac{K - N_t}{K},$$

where λ on the right-hand side takes on a different meaning (the maximum growth rate). This makes the original equation

$$N_{t+1} = \lambda N_t \frac{K - N_t}{K}.$$

We frequently define the variable N as a fraction of the carrying capacity, which is easily done by setting the carrying capacity equal to 1.0, a transformation that does not change the qualitative behavior of the equation and makes it easier to work with. Thus we have

$$N_{t+1} = \lambda N_t (1 - N_t). \quad (27)$$

This equation is usually referred to as the logistic map (*map* because it maps N_1 into N_2) or the logistic difference equation. It has some remarkable features that will be explored in more detail in chapter 4.

We add a short technical note here. The logistic map is *not* what you get when you integrate the logistic differential equation and then solve for N_{t+1} in terms of N_t , although the perceptive reader might be excused for thinking so because both equations are called logistic. The logistic map is derived directly from first principles (as above). Integrating the logistic differential equation gives a completely different time interval map.

It is worth noting that either of Bleasdale and Nelder's equations (equations 25 and 26) can be viewed as a form of a discrete map, much like the

logistic map, although with slightly different properties (recall exercise 1.15) (Watkinson 1980). If we think of yield as the number of organisms that will be found in the population in the next generation, this equation becomes equivalent to an iterative map (like the logistic map). In fact, the Bleasdale and Nelder formulation has been repeatedly invented in different contexts. For example, in fisheries biology both the Ricker (1954) map and the famous Beverton and Holt (1957) model have properties that are similar to the Bleasdale and Nelder formulation, as does the model proposed by Hassell (1975) in the context of a predator–prey model. Indeed, when $b = 1$, the Beverton and Holt model is identical to either of the forms of Bleasdale and Nelder, namely,

$$N_{t+1} = \frac{\lambda N_t}{(1 + \alpha N_t)}, \quad (28)$$

and the Hassell model is the same as the first form of the Bleasdale and Nelder model, namely,

$$N_{t+1} = \frac{\lambda N_t}{(1 + \alpha N_t)^b}. \quad (29)$$

The Ricker model has a similar overall form, although its formulation is exponential, namely,

$$N_{t+1} = N_t e^{r(1 - bN_t)}. \quad (30)$$

All have the property that when N_{t+1} is plotted against N_t a nonlinear form emerges with a quadratic-like hump, a fact that translates into some interesting and important dynamic generalizations that we discuss further in chapter 4.

EXERCISES

- 1.16** Using the exponential map and starting with 0.01 individuals, project the population 50 times with values of $\lambda = 1.0, 1.5, 2.0, 3.0$, and 3.5 . How does the time series change as a function of λ ?
 - 1.17** In exercise 1.4 successive values of population density were plotted against one another (i.e., N_{t+1} was plotted against N_t) for the exponential equation. Repeat that exercise for $\lambda = 1.5, 2.0, 3.0$, and 3.5 , and then modify the equation to be the logistic map (equation 27) and make the graphs again, for $\lambda = 1.5, 2.0, 3.0$, and 3.5 .
 - 1.18** Plot N_{t+1} against N_t for the Beverton and Holt model, the Hassell model, and the Ricker model (equations 28–30), and compare them. Let $\lambda = 1$, $\alpha = 1$, and $b = 2$. Once you get the three models on a spreadsheet, experiment with different values of b , λ , and α .
-

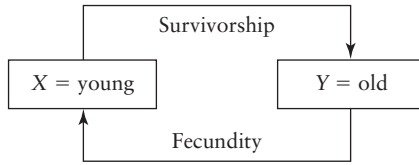
Elementary Age-Structured Population Projection Matrices

In the “unstructured” models discussed in chapter 1, we assume that all individuals are equal. Unstructured models are most often fitted to data on population sizes over time. Recent analyses of these sorts of data have become extremely sophisticated and will be discussed in a later chapter. However, most populations are divided into different classes of individuals. Insects have eggs, larvae, pupae, and adults. Plants have seeds, seedlings, saplings, and adults. Models of populations in which the individuals are thus “structured” are referred to as structured population models, the most common form of which are projection matrices. Structured models are often used to ask more detailed questions about populations, often relating to their management, their response to natural selection, or a variety of other specific issues.

Dividing individuals into classes can have dramatic effects, at least in the short term, on the overall dynamics of the population. A population of 10,000 eggs of some pest species has quite a different significance than a population of 10,000 adults, and a tree population with all seedlings is not usually called a forest, the name we would give to a tree population composed of mainly adult trees. Thus, taking account of age or stage classes complicates matters considerably.

The most general structure in populations is age. Populations are generally composed of individuals of different ages, and the unit used to classify the individuals is conveniently the same variable used to view the changes in the population, time. A child who is one year old this year will be two years old next year and eleven years old ten years from now.

The basic dynamics with two age classes, X and Y , can be visualized as follows:



The simple projection map for a population (as introduced in chapter 1) can be expanded to include various age classes. So rather than the simple

$$N_{t+1} = \lambda N_t,$$

where N stands for the whole population, we may have the two equations,

$$X_{t+1} = mY_t \quad (1a)$$

and

$$Y_{t+1} = pX_t, \quad (1b)$$

where X is the population density of the youngest age class and Y is the population density of the oldest age class (and $N = X + Y$, assuming that we can divide the population into just two age classes), m is the number of young individuals produced by each old individual (fecundity), and p is the probability that a young individual will survive to be an old individual (survivorship). Equations 1a and b can be conveniently represented as a matrix equation (readers who have forgotten their matrix algebra may consult the appendix to this chapter),

$$\begin{vmatrix} X_{t+1} \\ Y_{t+1} \end{vmatrix} = \begin{vmatrix} 0 & m \\ p & 0 \end{vmatrix} \begin{vmatrix} X_t \\ Y_t \end{vmatrix},$$

or in the more compact matrix form (following the customary procedure of printing vectors and matrices in boldface type)

$$N_{t+1} = \mathbf{P}N_t, \quad (2)$$

where

$$N_t = \begin{vmatrix} X_t \\ Y_t \end{vmatrix}$$

and

$$\mathbf{P}_t = \begin{vmatrix} 0 & m \\ p & 0 \end{vmatrix}.$$

Equation 2 is simply another way of writing equations 1a and 1b. The vector N is commonly referred to as the age distribution vector and the matrix $\mathbf{P}$ is called either the projection matrix or the Leslie matrix, named after the person who is generally credited with the development of the method (Leslie 1945), although the procedure had actually been published three years earlier (Lewis 1942).

Of course most populations that have stages or ages will have more than just two, but this way of representing the population is easily generalizable. For example, letting X_0 represent the population density of the youngest individuals, X_1 that of the one-year-olds, X_2 that of the two-year-olds, and so forth, up to X_n , the population density of the n -year-olds (assuming that n is the oldest the individuals in this population can become), equations 1a and 1b can be simply expanded as follows:

$$\begin{aligned} X_0(t+1) &= m_1 X_1(t) + m_2 X_2(t) + m_3 X_3(t) + \dots + m_n X_n(t), \\ X_1(t+1) &= p_{10} X_0(t), \\ X_2(t+1) &= p_{21} X_1(t), \\ X_3(t+1) &= p_{32} X_2(t), \\ &\vdots \\ X_n(t+1) &= p_{n,n-1} X_{n-1}(t), \end{aligned}$$

where the time variable is put in parentheses so as not to confuse it with the subscript that indicates the age category. This relatively large set of equations can be rewritten exactly as before, in the form of equation 2; here the age distribution vector includes all of the X s from X_0 to X_m , and the matrix P is

$$P = \begin{vmatrix} 0 & m_1 & m_2 & m_3 & \dots & m_n \\ p_{10} & 0 & 0 & 0 & \dots & 0 \\ 0 & p_{21} & 0 & 0 & \dots & 0 \\ 0 & 0 & p_{32} & 0 & \dots & 0 \\ \vdots & & & & & \\ 0 & \dots & & p_{n,n-1} & & 0 \end{vmatrix}.$$

The first subdiagonal contains the survivorship probabilities and the first row the fecundities. As in the previous case, we can write

$$N_{t+1} = PN_t,$$

which is identical to equation 2.

EXERCISES

2.1 Given the projection matrix

$$\begin{vmatrix} 0 & 1 & 4 \\ 0.7 & 0 & 0 \\ 0 & 0.2 & 0 \end{vmatrix}$$

begin with 10 individuals in the first age class and project the population 15 times (see the appendix). Plot the overall population over time.

- 2.2** From the data in exercise 2.1, plot the logarithm of the population density over time. What equation would best fit these data? Also graph the data from only the last five time periods (times 11–15). What equation would best fit these data?
- 2.3** From the projection in exercise 2.1, plot the population densities of each of the age categories over time. Compare the pattern with the results in exercise 2.1.
- 2.4** Given the projection matrix

$$\begin{vmatrix} 0 & 3 & 8 & 1 \\ 0.7 & 0 & 0 & 0 \\ 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \end{vmatrix}$$

suppose the population is introduced onto an island with 10 of the oldest individuals. Project the population 20 times, and graph the total number over time.

- 2.5** Using the projection matrix in exercise 2.4, suppose the population is introduced onto an island with 10 of the third age category (the next-to-oldest). Project the population 20 times, and graph the total number over time. Compare to the results of exercise 2.4.
- 2.6** Using the results from exercise 2.5, calculate the ratio of the number of individuals in category 2 to those in category 1, and plot that ratio over time, from time 3 to time 20. Do the same for categories 4 and 3. Compute the fraction of the total that is in the third category, and plot it over time.

Note the similarity between equation 2 (and the results of exercises 2.1–2.6) and the exponential map discussed in chapter 1. They are identical except that equation 2 deals with matrices and vectors. Indeed, the two equations are really two sides of the same coin, and the exponential map is simply a special case of equation 2 (the case with a single age category). However, an additional complication arises with the introduction of age categories into the model. Consider, for example, a population with three age categories, X_1 , X_2 , and X_3 . Suppose that the projection matrix is

$$P = \begin{vmatrix} 0 & 5 & 10 \\ 0.5 & 0 & 0 \\ 0 & 0.2 & 0 \end{vmatrix}$$

and the population begins with 10 individuals, all in category X_3 (for example, an island has just been invaded by 10 old individuals of the population). What will be the history of the population for the next few years? Applying equation 2, we obtain, for successive years,

$$\begin{vmatrix} 0 \\ 0 \\ 10 \end{vmatrix} \rightarrow \begin{vmatrix} 100 \\ 0 \\ 0 \end{vmatrix} \rightarrow \begin{vmatrix} 0 \\ 50 \\ 0 \end{vmatrix} \rightarrow \begin{vmatrix} 250 \\ 0 \\ 10 \end{vmatrix} \rightarrow \begin{vmatrix} 100 \\ 125 \\ 0 \end{vmatrix} \rightarrow \begin{vmatrix} 625 \\ 50 \\ 25 \end{vmatrix} \rightarrow \begin{vmatrix} 500 \\ 312 \\ 10 \end{vmatrix} \rightarrow \begin{vmatrix} 1,662 \\ 250 \\ 62 \end{vmatrix}.$$

The total population for those eight years will have been 10, 100, 50, 260, 225, 700, 822, and 1,975. The proportion of the total population in a particular age class varies enormously. For example, the oldest age class begins by representing 100% of the population (10 old individuals out of a total of 10 individuals), then 0%, then 0% again, then 4% (10/260), then 0%, then 4% again, then 1%, and finally 3%. If the projection is carried further, we find that this percentage stabilizes at around 2%. The same calculations could be done for the other age classes, making it clear that even though there is great initial variation in the proportion of the population represented by each of the age classes, after some time the proportions stabilize. In this particular example, the young age class stabilizes at about 76%, the second category at about 22%. So if we imagine an alternative example of initiating a population of 10 individuals in which the distribution of individuals is

$$\begin{vmatrix} 7.6 \\ 2.2 \\ 0.2 \end{vmatrix}$$

(granted, you can't have fractional individuals, but this is just an example), the future populations will look like this:

$$\begin{vmatrix} 7.6 \\ 2.2 \\ 0.2 \end{vmatrix} \rightarrow \begin{vmatrix} 13 \\ 4 \\ 0.4 \end{vmatrix} \rightarrow \begin{vmatrix} 23 \\ 6 \\ 0.76 \end{vmatrix} \rightarrow \begin{vmatrix} 40 \\ 12 \\ 1 \end{vmatrix} \rightarrow \begin{vmatrix} 72 \\ 20 \\ 2 \end{vmatrix} \rightarrow \begin{vmatrix} 124 \\ 36 \\ 4 \end{vmatrix} \rightarrow \begin{vmatrix} 219 \\ 62 \\ 7 \end{vmatrix} \rightarrow \begin{vmatrix} 381 \\ 109 \\ 12 \end{vmatrix}.$$

We see that each age category is increasing, but if you calculate the percentage representation of each age category in the population as a whole, it remains perfectly constant or, loosely speaking, stable. When the distribution of individuals in a population has this form, it is referred to as a “stable age distribution.”

EXERCISES

2.7 Consider the following projection matrix representing a population with five age classes:

$$\begin{vmatrix} 0 & 5 & 3 & 2 & 1 \\ 0.9 & 0 & 0 & 0 & 0 \\ 0 & 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0 & 0.05 & 0 \end{vmatrix}$$

Begin with a population distributed as 0, 0, 0, 0, and 10 individuals, project the population 20 time units, and plot the total number over time. Plot the natural log of the numbers over time.

2.8 Repeat exercise 2.7, but begin with 80, 16, 5, 1, and 1 individuals. Compare the results with the results from exercise 2.7.

Continuing with this same example, when the population is at a stable age distribution, because each age category is increasing by the same proportional amount, each of the separate equations must have the same constant rate of increase, which is to say

$$X_i(t+1) = \lambda X_i(t),$$

for all values of i , where we use the constant λ to emphasize that the rate of increase of each category is the same as the finite rate of increase, as discussed in chapter 1, here referring to that rate at which each of the population density categories is increasing (also the rate at which the entire population is increasing). In matrix notation we have

$$N_{t+1} = \lambda N_t. \quad (3)$$

Furthermore, if N (not boldfaced) is equal to the total population, we have

$$N_{t+1} = \lambda N_t,$$

which is identical to the exponential map (see chapter 1). In figure 2.1, the above examples are plotted over time. It is evident that the population beginning with a stable age distribution increases in a smooth exponential fashion, as normally expected, while the population beginning with 10 individuals in the oldest category has significant fluctuations before increasing in a smooth, typically exponential, fashion.

It is a well-known rule that any population with constant mortality and natality will eventually reach a stable age distribution, at which point it will be growing according to the classic exponential equation. (There are technical

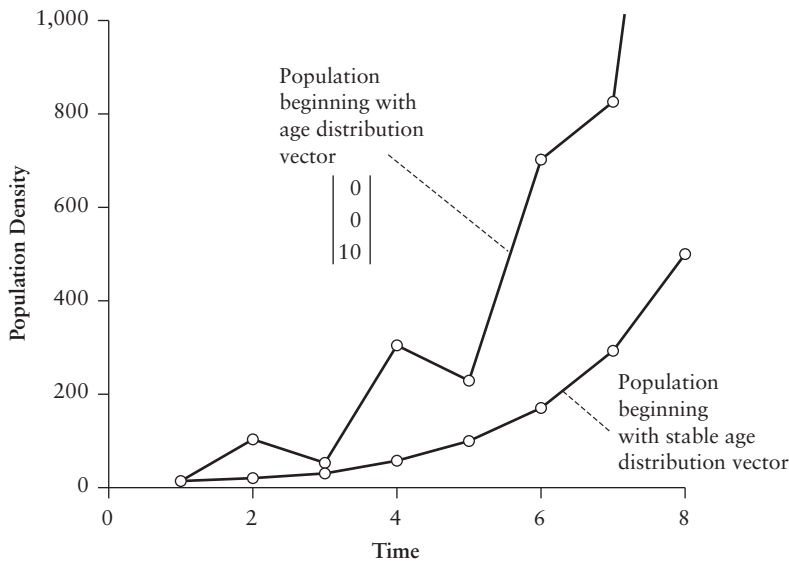


FIGURE 2.1. Population trajectories for different starting points in an age-distributed population.

exceptions to this rule, but they need not concern us presently. See Caswell 2001 for a more detailed discussion.) As mentioned previously, most populations do not grow according to the exponential equation for long; some form of density dependence usually sets in, or some controlling factor (a parasite or predator, for example) limits the population. Introducing density dependence in an age- or stage-distributed model is a relatively complicated affair and will be discussed later in this chapter.

Any structured population at its stable age distribution increases in an exponential fashion. We can quite easily take such a population and simply project it into the future and calculate the rate at which it is growing. That rate will be the same as we have seen before, the intrinsic rate of natural increase. But there is another way of calculating the rate of increase directly from the matrix. Mathematically speaking, the finite rate of increase is the same as the dominant eigenvalue of the projection matrix, as discussed below.

EXERCISE

- 2.9** Using the projection matrix from exercise 2.4, begin with a population vector of 674, 263, 47, 3 (the stable age distribution), and project the population 10 time units. Plot the natural log of the total population density against time. Then calculate the intrinsic rate of natural increase from this graph.
-

As noted earlier, a population with n age categories will have a matrix that looks like this:

$$P = \begin{vmatrix} 0 & m_1 & m_2 & m_3 & \dots & m_n \\ p_{10} & 0 & 0 & 0 & \dots & 0 \\ 0 & p_{21} & 0 & 0 & \dots & 0 \\ 0 & 0 & p_{32} & 0 & \dots & 0 \\ \vdots & & & & & \\ \vdots & & & & & \\ 0 & \dots & & p_{n,n-1} & & 0 \end{vmatrix}.$$

and the basic process of population projection remains

$$N_{t+1} = PN_t. \quad (4)$$

We can also write

$$N_{t+2} = PN_{t+1}. \quad (5)$$

Substituting equation 4 into equation 5, we get

$$N_{t+2} = PPN_t = P^2N_t,$$

and in general

$$N_{t+n} = P^n N_t. \quad (6)$$

Repeating the calculations already made in chapter 1, take equation 3,

$$N_{t+1} = \lambda N_t,$$

and note that

$$N_{t+2} = \lambda N_{t+1}.$$

Repeated substitution yields the generalization

$$N_{t+n} = \lambda^n N_t. \quad (7)$$

Note that equations 6 and 7 can be combined to give

$$P^n N_t = \lambda^n N_t, \quad (8)$$

where we see the remarkable result that we can substitute a simple scalar for the entire projection matrix (i.e., λ^n is not a matrix but an ordinary scalar number) when the population is at a stable age distribution. Normally the value of λ is set equal to e^r , where r is then the intrinsic rate of natural increase (equivalent to r of the previous chapter).

EXERCISES

2.10 Using the same projection matrix as in exercise 2.9, multiply the matrix by itself 10 times. Recall that you estimated the intrinsic rate of natural increase as 0.5603 in exercise 2.9. Verify equation 8 by calculating the vector N_{t+1} , first from $P^{10}N_t$ and then from $\lambda^{10}N_t$, using the stable age distribution vector for N_t and then plotting the two calculations against one another.

2.11 Repeat exercise 2.10 for the population vector $\begin{vmatrix} 200 \\ 300 \\ 50 \\ 40 \end{vmatrix}$, which of course is not the stable age distribution.

If we now set the beginning of the projection as $n = 1$, we have

$$N_{t+1} = \lambda N_t, \quad (9)$$

which is a fundamental property of age-distributed populations. That is, as long as t is relatively large, equation 9 says that each time projection results in each age category multiplied by a constant, and consequently the whole population is also increased by that amount. This means that λ must be identical to e^r , as first discussed in chapter 1.

We are now prepared to go one step further and derive a most important equation in matrix population models, the one that enables us to directly calculate the intrinsic rate of natural increase (frequently referred to as the dominant eigenvalue in population models; more on eigenvalues in chapter 4). Recall equation 8, and set $n = 1$ to obtain the following equation:

$$PN_t = \lambda N_t.$$

So as to be able to subtract two matrices (which we need to do to arrive at equation 11), we convert the right-hand side to a square matrix,

$$PN_t = \lambda IN_t, \quad (10)$$

where I is the identity matrix—a square matrix with 1s on the principal diagonal and 0s everywhere else. (Again, if your matrix algebra is rusty, refer to the appendix).

Equation 10 can be manipulated so as to give

$$(P - \lambda I)N_t = 0. \quad (11)$$

Consider equation 11 for an annual plant population that has only two age categories (recall equation 1), adults (who produce m juveniles in each time unit but do not survive more than a single year) and juveniles (who have a survival rate of g). The projection matrix will be

$$\begin{bmatrix} 0 & m \\ g & 0 \end{bmatrix}.$$

Equation 11, for this example, gives us

$$\left[\begin{bmatrix} 0 & m \\ g & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right] \begin{bmatrix} N_0 \\ N_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or

$$\begin{bmatrix} -\lambda & m \\ g & -\lambda \end{bmatrix} \begin{bmatrix} N_0 \\ N_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Multiplying the matrix equation, we obtain

$$-\lambda N_0 + mN_1 = 0 \quad (12a)$$

and

$$gN_0 - \lambda N_1 = 0. \quad (12b)$$

From equation 12a we write

$$N_0 = \frac{mN_1}{\lambda}$$

and, substituting into equation 12b, we have

$$\frac{gmN_1}{\lambda} = \lambda N_1,$$

which simplifies to

$$(\lambda^2 - gm)N_1 = 0,$$

which is true if and only if (presuming that $N_1 > 0$)

$$\lambda^2 - gm = 0.$$

Note that the quantity on the left-hand side is, by definition, the determinant of

$$\begin{vmatrix} -\lambda & m \\ g & -\lambda \end{vmatrix},$$

which gives us the very basic equation

$$\text{Det}(\mathbf{P} - \lambda \mathbf{I}) = 0.$$

This equation is referred to as the characteristic equation of the matrix $\mathbf{P}$ and is the equation used to find the value of λ when $\mathbf{P}$ is known (actually there are two values here, but for an n -dimensional system there will be n values). In the context of a population projection matrix, the largest of those roots is the rate of growth of the population after it has been projected a number of times with the same projection matrix and has reached a stable age distribution.

EXERCISES

2.12 Given the projection matrix $\begin{vmatrix} 0 & m \\ g & 0 \end{vmatrix}$, find the two eigenvalues.

2.13 Given the projection matrix $\begin{vmatrix} 0 & 3 \\ 0.9 & 0 \end{vmatrix}$, compute the eigenvalues. Beginning with $N_1 = 21.5$ and $N_2 = 23.6$, project the matrix forward 30 times, plot the log of the total number against time, and estimate the rate of population growth from a line drawn on the plot (recall from the previous chapter what the exponential rate of increase looks like after taking natural logs). Does the estimate of the finite rate of population increase correspond to the computation of the dominant eigenvalue? Note here that a model in which only one category produces offspring has some unusual behaviors, and we include one here only to illustrate the general correspondence with the direct calculation of the eigenvalues and the projected population.

2.14 Given the projection matrix $\begin{vmatrix} 0 & 3 & 8 & 1 \\ 0.7 & 0 & 0 & 0 \\ 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \end{vmatrix}$, compute the eigenvalues.

Project the population starting with the population vector $\begin{vmatrix} 76 \\ 86 \\ 99 \\ 14 \end{vmatrix}$, and directly estimate

the dominant eigenvalue. (Note that in these exercises the population behaves with an apparent cycling of either 3 (exercise 2.13) or 4 (exercise 2.14) time units. This is because there is only a single category that produces offspring, a restriction we put on the exercise so the eigenvalue could be computed analytically).

Non-Age Structure: Stage Projection Matrices

In many situations the use of a Leslie–Lewis matrix is very inconvenient. Consider, for example, a population of long-lived trees, say oak trees that live as long as 200 years. We could define ages as quarter centuries (25 years) and have an eight-by-eight projection matrix. But a great deal of interest actually hap-

pens at the seed and seedling stages, which are lumped into the first 25 years as if they were part of the same age category (a seed, a seedling, and a 25-year-old sapling can hardly be considered equivalent). On the other hand, if we define the ages as one-year intervals (which would make sense for the seeds, which may live for a maximum of a single year), we are stuck with a 200-by-200 projection matrix, an unwieldy object to analyze, to say nothing of the prospect of having to estimate the survival and fecundity of individuals that differ from one another by a single year. A 100-year-old tree that has been growing in the shade and reaches only 10 meters in height probably has survivorship and fecundity properties similar to those of a 30-year-old tree that has been growing in the sun and reaches that same 10 meters. That is, for fecundity and survivorship, size may be a much better predictor than age. Similarly, for insects with discrete life-cycle stages such as eggs, larvae, pupae, and adults, stage rather than age may be much more important in determining vital rates.

When populations are structured by stages (including size classes) rather than age, models do not retain the characteristics of the original Leslie–Lewis formulation. Nevertheless, we often need to analyze such stage-structured models. For example, a particular species of tropical tree is sought for its lumber. It is, of course, impossible to do an experiment to decide what harvesting regime allows for a sustainable yield of timber (i.e., how much can be harvested while retaining a viable population in perpetuity) because such an experiment would take hundreds of years to complete. The only option is to develop a model of the population. Because it is a highly structured population (seeds are different from seedlings, which are different from saplings, etc.), the most obvious model is one based on stage classes.

In some ways the need for stages rather than ages is unfortunate, because the age-structured projection matrix has a very convenient structure (all zeros except the first subdiagonal and the first row). But in many other ways it is the form of the projection matrix equation that matters, and the details of the structure of the projection matrix are irrelevant. Thus, much of the mathematical machinery is just as relevant for projection matrices based on stages as it is for matrices based on ages. The technique of using stage projection matrices was pioneered by Lefkovitch (1965), and stage-structured matrix models are sometimes referred to as Lefkovitch matrices.

EXERCISES

- 2.15** In a population of insects we find that there are, at a particular point in time, 1,000 larvae, 50 pupae, and 5 adults. We find that after another census one month later we are able to estimate the probability of survival of a larva (assuming that it remains a larva) as 0.7, the probability that the larva will turn into a pupa as 0.2, the probability that a pupa will survive as 0.1, the probability that it will turn into an adult as 0.4, and the probability that an adult will survive as 0.5. Furthermore, an average adult produces 100 eggs, 50% of which appear (within that month) as larvae while the rest die. Construct the projection matrix that takes into account all of these probabilities, and project the population one time unit (remember that the time unit is one month).

- 2.16** Project the population of exercise 2.15 ten times, and graph the total population over time. What is the rate of population increase?
- 2.17** Plot the proportion of each stage over time (in three different graphs).

Suppose that the population is conveniently divided into four stages: seed, seedling, sapling, and adult. How do we construct a stage projection matrix? The identification of the stages is obvious for the seeds, but we must provide an arbitrary measure that distinguishes the other three categories. For example, seedlings are all plants less than 1 meter tall, saplings are all plants between 1 meter tall and 10 centimeters in breast-height diameter, and adults are all larger plants. Suppose that we take a census of the population and discover that in one hectare there are 10,000 seeds, 5,000 seedlings, 100 saplings, and 50 adults. Thus, if we construct a population vector much as we did in the case of an age distribution vector, we have

$$\begin{pmatrix} 10,000 \\ 5,000 \\ 100 \\ 50 \end{pmatrix}$$

Now suppose that we are able to mark all individuals in the population (even all the seeds—although what we would undoubtedly do in practice is mark a subsample), and then we return next year to sample again. We find that 9 out of 10 seeds died and the other 10% germinated, thus making the probability of germination equal to 0.1. We find that of the 5,000 seedlings marked, 1,000 are still seedlings, 5 became saplings (grew to a height of more than 1 meter), and the rest died. Thus the probability of a seedling's staying a seedling is $1,000/5,000 = 0.2$, and the probability of its growing to become a sapling is $5/5,000 = 0.001$. Of the saplings marked, we find that 30 died (probability of mortality = $30/100 = 0.3$), 2 became adults (grew to the point at which their breast-height diameter was greater than 10 centimeters) (probability of growing to adulthood = $2/100 = 0.02$), and the rest (68) remained as saplings (probability of remaining as saplings = $68/100 = 0.68$). Five of the adults fell over (probability of surviving as an adult = $45/50 = 0.9$). Finally, we make a count of all the seeds (all of which must be new, because seeds last only a year) and discover 10,000 (which means that, on average, each of the 10 adult trees produced 1,000 seeds). With these data we can write a series of four equations for the projection of each of the stage classes, letting N_0 = number of seeds, N_1 = number of seedlings, N_2 = number of saplings, and N_3 = number of adults. Namely,

$$N_0(t+1) = 1,000 N_3(t), \quad (13a)$$

$$N_1(t+1) = 0.1 N_0(t) + 0.2 N_1(t), \quad (13b)$$

$$N_2(t+1) = 0.001 N_1(t) + 0.68 N_2(t), \quad (13c)$$

and

$$N_3(t + 1) = 0.2 N_2(t) + 0.9 N_3(t). \quad (13d)$$

Solving these four equations, we find that the population of the second year is easily computed from the population of the first year. Furthermore, equations 13a–13d can be rewritten in matrix form as

$$\begin{bmatrix} N_0(t + 1) \\ N_1(t + 1) \\ N_2(t + 1) \\ N_3(t + 1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1,000 \\ 0.1 & 0.2 & 0 & 0 \\ 0 & 0.001 & 0.68 & 0 \\ 0 & 0 & 0.2 & 0.9 \end{bmatrix} \begin{bmatrix} N_0(t) \\ N_1(t) \\ N_2(t) \\ N_3(t) \end{bmatrix}$$

or

$$\mathbf{N}_{t+1} = \mathbf{P}\mathbf{N}_t,$$

where, as before, boldface indicates matrices and vectors. Once again the projection matrix $\mathbf{P}$ is a square matrix, but this time there are more entries than the simple subdiagonal and first row of the classic age projection matrix. If we repeatedly apply the projection matrix to the initial stage distribution vector for five years, we obtain the series of vectors as follows:

$$\begin{bmatrix} 10,000 \\ 5,000 \\ 100 \\ 50 \end{bmatrix} \rightarrow \begin{bmatrix} 50,000 \\ 2,000 \\ 73 \\ 65 \end{bmatrix} \rightarrow \begin{bmatrix} 65,000 \\ 5,400 \\ 52 \\ 73 \end{bmatrix} \rightarrow \begin{bmatrix} 73,100 \\ 7,580 \\ 40 \\ 76 \end{bmatrix} \rightarrow \begin{bmatrix} 76,118 \\ 8,826 \\ 35 \\ 77 \end{bmatrix} \rightarrow \begin{bmatrix} 76,609 \\ 9,377 \\ 33 \\ 76 \end{bmatrix} \rightarrow \dots$$

The population appears to be growing, and, especially if you follow the adult trees only (which an ecosystem manager or forester is likely to do), you could come to the conclusion that the population is indeed in a healthy state, perhaps stabilizing at about 76 adult trees per hectare. But if we apply the matrix many more times, after 200 years we get the pattern shown in figure 2.2 (light line). Clearly this was not a viable population. Although it may take 200 years, with the vital statistics in matrix $\mathbf{P}$, we can expect that the population will be effectively extinct after a couple of centuries.

This example also provides an interesting exercise in the use of such models in ecosystem management. Based on the pattern seen after five years of observation of this population (assuming that the numbers above are actual numbers of individuals in a real population that can be observed), the forester may conclude that not only is this a healthy population, but it could be beneficially harvested. Because it changed from only 50 trees per hectare to more than 75 trees per hectare in only five years, one might conclude that harvesting some of the adult trees certainly could not have a big effect on the population. So the forester might conclude that a 5% harvest of the standing stock (adult trees) would certainly not hurt the population (i.e., with 50 trees per hectare, only 2.5 trees per year would be harvested, and with 75 trees per

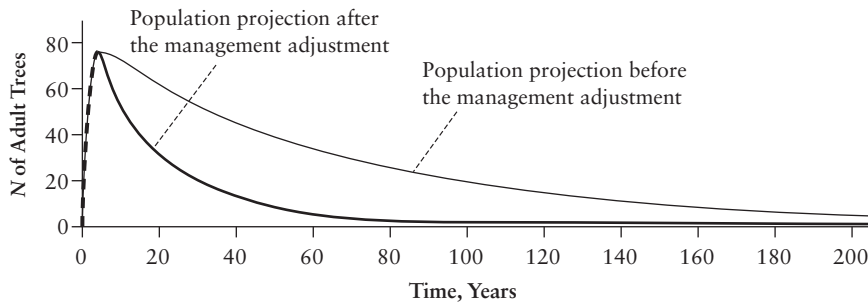


FIGURE 2.2. Long-term projections of the hypothetical tree population discussed in the text. Before the management adjustment, for the first 5 years the population appears to be growing, but then it begins a rapid decline to become very rare after 200 years. After changing the harvesting regime (boldface trajectory), based on observations of the initial growth period (the first 5 years), the population declines even more rapidly, becoming virtually extinct after only 100 years.

hectare, only 3.75 trees would be harvested per year). It is easy to see how one might think of such a minimal intervention as not significant for the overall health of the population. But changing the element p_{33} in the P matrix from 0.9 to 0.85 after the first five years and then projecting further population growth, we obtain

$$\begin{array}{c|c} \begin{array}{c} 76,609 \\ 9,377 \\ 33 \\ 76 \end{array} & \rightarrow \begin{array}{c} 76,000 \\ 9.536 \\ 32 \\ 71 \end{array} & \rightarrow \begin{array}{c} 71,200 \\ 9.507 \\ 31 \\ 67 \end{array} & \rightarrow \begin{array}{c} 66,883 \\ 9,021 \\ 31 \\ 63 \end{array} & \rightarrow \begin{array}{c} 63,085 \\ 8,493 \\ 30 \\ 60 \end{array} & \rightarrow \begin{array}{c} 59,763 \\ 8,007 \\ 29 \\ 57 \end{array} & \rightarrow \begin{array}{c} 56,778 \\ 7,577 \\ 28 \\ 54 \end{array} \end{array}.$$

Applying the P matrix (with the 0.9 probability changed to 0.85), we observe the pattern shown in figure 2.2B. After just six more years the adult population is almost back to where it was when the original observations were made (about 50 individuals), and the population is virtually extinct after only about 100 years. The harvesting intervention has a large effect indeed, and one that would not have been easy to see without the projection model.

The stage projection matrix is obviously a far more general model than the original age projection model. Yet the basic rules of matrix manipulation apply to stage projection matrices as well. Specifically if we have

$$N_{t+1} = PN_t, \quad (14)$$

we can also write

$$N_t = P_t N_0.$$

Any square matrix multiplied by itself a large number of times can be represented by a matrix with its dominant eigenvalue raised to the power of t on its principle diagonal and zeros everywhere else, which means we can also write

$$N_t = \lambda_t N_0,$$

which also means that

$$N_{t+1} = \lambda N_t, \quad (15)$$

which will be true only after the population has settled down to a stable stage distribution (the same as a stable age distribution except that here we are referring to stages rather than ages). Because the left-hand sides of equations 14 and 15 are identical, we can equate the right-hand sides to obtain

$$PN_t = \lambda N_t,$$

which can also be written

$$PN_t - \lambda N_t = 0$$

or

$$(P - \lambda I)N_t = 0,$$

which, as explained in a previous section, is true only when

$$\text{Det}(P - \lambda I) = 0. \quad (16)$$

As explained before, this is the characteristic equation that is used to find the eigenvalues of the projection matrix.

EXERCISES

- 2.18** In a population of plants we categorized individuals as small seedlings, large seedlings, small saplings, large saplings, and adults. Presuming that the plants in each stage either move to the next stage or stay in their own category during population growth, write the stage projection matrix in general terms.
- 2.19** At another site this same population has slightly different growth patterns and the variability of its growth is very large, such that occasionally the small seedlings grow all the way to small saplings during a single time unit. Furthermore, because of severe weather, occasionally a large sapling has its top damaged and actually reverts to the small sapling stage. Modify the matrix of exercise 2.18 to account for these probabilities.

It is important to realize that when populations are modeled as stages (these models are frequently referred to as *i*-stage models), there is really no restriction on where the transition probabilities must appear in the matrix. That is, in the strictly age-structured case all transition probabilities except the top row of fecundities are survivorship probabilities, and all appear precisely in the first subdiagonal of the matrix. But in the case of other structures defining the stage (size, morphological distinctness, life cycle stage, etc.), depending on the actual biology involved, probabilities can occur in almost any part of the matrix. However, there is a certain pattern to interpretations of transition probabilities in different parts of the matrix. If the stages are ordered in some biologically interesting way (e.g., small trees to large trees, early instars

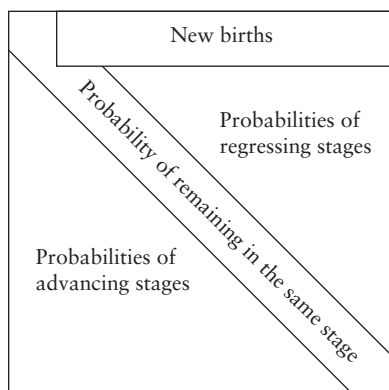


FIGURE 2.3. Diagrammatic summary of the general structure of a stage projection matrix.

to late instars), the subdiagonals refer to the probability of advancing stages and the supradiagonals refer to the probability of regressing, while the principal diagonal contains the probabilities of remaining in a stage (see figure 2.3). It may seem strange for individuals to “regress” stages, but such behavior can be common, depending on how stages are defined. For example, a tree population is usually modeled with size as the categorization variable. It is quite common, especially among seedlings and young saplings, for that size to sometimes decrease—for deer to browse on oak seedlings or large leaves to fall and cover tropical tree seedlings, killing their upper branches but allowing their lower branches to survive and form new leaves. Many such examples could be cited.

Eigenvectors, Reproductive Value, Sensitivity, and Elasticity

Recall the basic equation, now repeated at least twice in this chapter,

$$N_{t+1} = \lambda N_t.$$

As discussed previously, when a population follows this equation we refer to its age distribution as the stable age distribution, meaning that the proportion of individuals in each age category does not change as the projection matrix continues to be applied. The related equation

$$PN_t = \lambda N_t \quad (17)$$

basically says the same thing. When equation 17 is true, the column vector N is referred to as an eigenvector (recall that λ is the eigenvalue), and we write it as

$$Pu = \lambda u, \quad (18)$$

using u as a special symbol for the special column vector that satisfies equation 17. The vector u is thus the stable stage (or age) distribution vector. It is also the case that if we take the transpose of P (exchange the rows and columns of the matrix) and symbolize the transpose as P^* , we can write

$$P^*v = \lambda v, \quad (19)$$

where the symbol v is used as a special symbol for the vector that satisfies equation 18. Exactly what is the meaning of this vector?

According the formal terminology of matrix algebra, when a vector satisfies equation 17, the parameter λ is the eigenvalue. The vector u is referred to as the right eigenvector and the vector v as the left eigenvector (the left eigenvector because P^*v can also be written as v^*P , where the asterisk refers to the transpose of the matrix— v^* is a row matrix, while v is a column matrix—and thus v^* is multiplied on the left, while u is multiplied on the right by the matrix P). Taking a simple example, we write equation 18 as

$$\begin{vmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{vmatrix} \begin{vmatrix} u_1 \\ u_2 \\ u_3 \end{vmatrix} = \begin{vmatrix} \lambda u_1 \\ \lambda u_2 \\ \lambda u_3 \end{vmatrix} \quad (20)$$

and equation 19 would be

$$\begin{vmatrix} p_{11} & p_{21} & p_{31} \\ p_{12} & p_{22} & p_{32} \\ p_{13} & p_{23} & p_{33} \end{vmatrix} \begin{vmatrix} v_1 \\ v_2 \\ v_3 \end{vmatrix} = \begin{vmatrix} \lambda v_1 \\ \lambda v_2 \\ \lambda v_3 \end{vmatrix}, \quad (21)$$

where p_{ij} is the probability of transitioning from stage j to stage i . The difference between equations 20 and 21 is that the square matrix of 21 is the transpose of the square matrix of 20. We already know that the right eigenvector contains the elements of the stable age distribution. What might be the meaning of the left eigenvector, and might it have some sort of biological significance?

To answer this question, first write out one of the linear equations of the matrix equation 20, say the second equation. We have

$$p_{21}u_1 + p_{22}u_2 + p_{23}u_3 = \lambda u_2,$$

which can also be written as

$$u_2 = (p_{21}u_1 + p_{23}u_3)/(\lambda - p_{22}). \quad (22)$$

Note that on the left-hand side is the single element referring to the value of the second stage in the stable stage distribution, while on the right-hand side are all the probabilities that contribute to that stage (the first subscript on all the probabilities is 2). Now do the same multiplication for equation 21, obtaining

$$p_{12}v_1 + p_{22}v_2 + p_{32}v_3 = \lambda v_2,$$

which can also be written as

$$v_2 = (p_{12}v_1 + p_{32}v_3)/(\lambda - p_{22}). \quad (23)$$

Here we have on the left-hand side the single element referring to the value of the second stage in the left eigenvector, while on the right-hand side are all the probabilities to which that stage contributes. In other words, v_2 represents the

total contribution of the second stage category to the other categories (the second subscript on all the probabilities is 2). For this reason, the elements of the left eigenvector are called the reproductive value (not an especially good choice of term but one we seem to be stuck with from extensive use in the literature). As is evident from the difference between equations 22 and 23, the elements of the right eigenvector refer to the contribution of all other stages to that stage (the stable stage distribution), while the elements of the left eigenvector refer to the contribution of that stage to all other stages (the reproductive value).

Once a projection matrix has been estimated from real data, the basic information of interest to the population ecologist is the rate of population increase (the dominant eigenvalue, λ), the stable age distribution (the right eigenvector), and the distribution of reproductive values (the left eigenvector). But although these figures may be of interest in and of themselves, it is frequently even more interesting to know how they might change under different circumstances. Such a desire may be in the mind of an ecosystem manager who wishes to set size limits on a sport fishery or extraction limits on a nontimber forest product in a rain forest or many other possible examples. Or an evolutionary biologist might be interested in knowing how fitness (i.e., per capita population growth or λ) would change with a change in life history, such as an earlier or later age of first reproduction (see chapter 3). We wish to know how λ changes when we make changes in p_{ij} . Mathematically we seek the value of

$$s_{ij} = \partial\lambda/\partial p_{ij},$$

where s_{ij} is referred to as the “sensitivity” of the matrix with respect to p_{ij} . The meaning of sensitivity is clear. How much will the rate of population growth change with a change in p_{ij} ? The value of s_{ij} is given by the simple formula (see Caswell 1978 or 2001 for a derivation)

$$s_{ij} = v_i u_j, \quad (24)$$

where v_i and u_j are the appropriate elements of the left and right eigenvectors (the reproductive value of the i th stage times the stable stage element of the j th stage). As elegant a formula as is equation 24, its utility is limited because of scaling problems. For example, transition probabilities are always between 0 and 1, while the birth rate is frequently very large (thousands for some plants, millions for corals, etc.). So the change in λ from a “unit” change in p_{ij} can be very misleading. For this reason, it has been convenient to develop a “proportional” representation of the same idea, a measure that would give a fractional contribution to λ of a change in p_{ij} . Caswell et al. (1984) proposed the use of a concept borrowed from microeconomics, the “elasticity” of a parameter. The elasticity is given as the rate of change in the log of λ with respect to the log of an element of P . Specifically,

$$e_{ij} = \partial(\ln\lambda)/\partial(\ln p_{ij}), \quad (25)$$

which (since $d\ln x = (1/x)dx$) is

$$e_{ij} = \frac{p_{ij}}{\lambda} \frac{\partial\lambda}{\partial p_{ij}} = \frac{p_{ij}}{\lambda} v_i u_j. \quad (26)$$

which is the formula that can be used for calculating the elasticities (find the dominant eigenvalue and the two eigenvectors, and do the appropriate multiplication). It is not at all obvious, given the defining statement for elasticity (equation 25), but the sum of all elasticities in the whole matrix is equal to 1.0 (de Kroon et al. 1986). This fact means that we can interpret e_{ij} as the proportional sensitivity of λ to changes in p_{ij} (because all the e_{ij} s sum to 1.0, that makes them interpretable as proportions). However, it is frequently useful to look at both sensitivity and elasticity when interpreting a population. Sensitivity is especially useful when entries in the matrix are zero. With a zero entry, the elasticity is automatically 0 (see equation 25). This means that it is not possible to assess the importance of elements that are zero, even though a small change in one of them might result in a large change in λ (this problem is due to the original logarithmic definition of elasticity, equation 25).

Density Dependence in Structured Populations

Thus far all of the material in this chapter has been developed under the assumption that vital statistics are not dependent on the density of the population. We have tacitly assumed that the forces acting on an oak seedling are the same in an open field as in the understory of an oak forest, that a lone coral hydroid on a cement block in the middle of the ocean has a similar survivorship as one in the middle of the Great Barrier Reef, that a bacterial cell isolated in the middle of a nutrient agar petri dish is equivalent to a bacterial cell in the middle of a petri dish that is spilling over with bacteria. All seem to be ridiculous assumptions yet might be reasonable under certain circumstances. Most generally, empirical measurements of transition probabilities will reflect the influence of current densities. As long as density doesn't change too much, estimating the current rate of population growth (or decline) for conservation purposes using a project matrix without explicit density dependence could be quite useful. The density-independent assumption is perfectly appropriate for the short term, although it would be folly to attempt a long-term projection. In some cases populations do indeed behave as if they were density independent, at least for short periods of time, and modeling them over those short periods of time is reasonable.

However, the general consensus is that if we wish to understand the behavior of a population over a long period of time, we must modify the density-independent assumption. Just as we had to modify the exponential equation to take into account the obvious forces of density dependence, we must do the same for structured populations. Unfortunately, things become very complicated very quickly.

Density Dependence in a Simple Age-Structured Model

Consider a population in two distinct phases, say, larva (X) and adult (Y). Presuming both larva and adult live exactly one time period (one time period is necessary for larval development and metamorphosis, and all adults die after one time period), the projection matrix would be as discussed earlier,

$$\begin{vmatrix} X_{t+1} \\ Y_{t+1} \end{vmatrix} = \begin{vmatrix} 0 & m \\ p & 0 \end{vmatrix} \begin{vmatrix} X_t \\ Y_t \end{vmatrix}. \quad (27)$$

As explained above, a projection matrix with constant coefficients always represents an exponentially growing population, yet most populations do not grow exponentially. To incorporate density-dependent effects, much as we developed the logistic equation in the previous chapter, we expand the model (equation 27) to allow for nonlinearities (density-dependent effects) in the fecundity factor (m)—certainly not the only possibility but the example we arbitrarily choose here. Assume that the fecundity factor is

$$m = r \left(1 - \frac{Y_t}{r} \right),$$

where we presume an upper limit of r on the rate of production of offspring and presume that there is an inverse relationship between the maximal rate of offspring production and the density-dependent force (i.e., if the potential rate of reproduction is very high, the feedback from density is low). Equation 27 thus becomes

$$\begin{vmatrix} X_{t+1} \\ Y_{t+1} \end{vmatrix} = \begin{vmatrix} 0 & r \left(1 - \frac{Y_t}{r} \right) \\ p & 0 \end{vmatrix} \begin{vmatrix} X_t \\ Y_t \end{vmatrix}.$$

Multiplying this matrix equation, we obtain the two-dimensional map (two-dimensional because we have two state variables, X and Y)

$$X_{t+1} = rY_t \left(1 - \frac{Y_t}{r} \right), \quad (28a)$$

and

$$Y_{t+1} = pX_t. \quad (28b)$$

If we begin with equation 27 and then plot the total population ($X + Y$) over time, we get, as expected, an exponentially growing population. When we add density dependence (i.e., equation set 28), the population tends to stabilize, as shown in figure 2.4. However, that stabilization has a curious aspect to it, a product of this sort of model, to be sure, but also a natural consequence of adding nonlinear elements to the system. Thus in figure 2.4 we see first the exponential pattern when there is no density effect. Adding density dependence does indeed make the population behave, qualitatively, like the logistic equation. Yet the nature of the pattern depends on the magnitude of the intrinsic rate of offspring production, as is clear in figure 2.4A. Furthermore, if the offspring production becomes very large, the overall pattern becomes virtually indistinguishable from randomness, what is now routinely called chaos (figure 2.4B). And this is with only two age categories. If a similar procedure were applied to populations with more age categories, even with very simple nonlinear functions modeling the density-dependent effects in each age

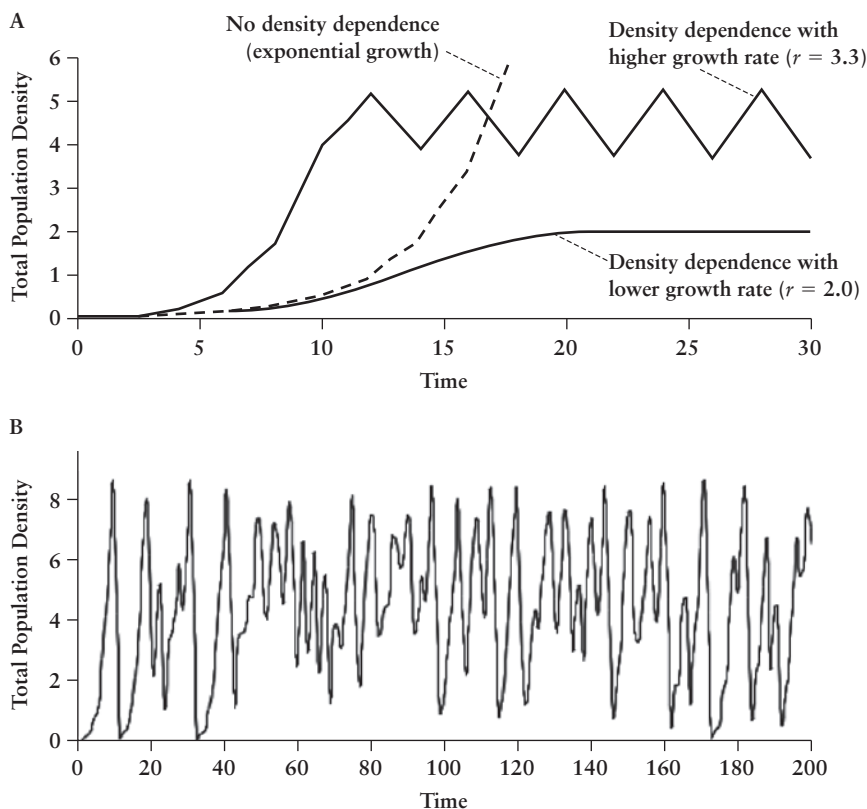


FIGURE 2.4. Time series resulting from density dependence in a simple age-structured model (equations 28a and 28b). (A) Illustrations of no density dependence (dashed line) and two levels of intrinsic rate of reproduction ($r = 2$ and $r = 3.3$) that lead to qualitatively different dynamics when combined with density dependence. (B) Unpredictable chaotic time series resulting from the same model with very large density dependence ($r = 4.2$). $p = 0.95$ for all trajectories.

category, the resulting population projections might be even more complicated. The possibilities for extremely complicated behavior of the population through time are enormous. In chapter 4 we revisit these ideas in more detail.

Density Dependence in Size-Distributed Populations

The above discussion was made relatively simple by the assumption that we were dealing with an age-distributed population. We now turn to the question of size-distributed populations.

We focus on size because size is clearly important for the fate of individuals. As far as we know, the largest individual organism living in the world today is not a blue whale, as popularly believed, but an individual of the fungus *Armillaria gallica*, estimated to weigh more than 100 tons (also estimated to be 1,500 years old). And this individual began as a single cell. The differ-

ence between a single cell and a 100-ton individual can hardly be ignored. Small individuals are relatively more likely to die due to inability to secure resources or failure to withstand physical forces. Predation or herbivory is often size dependent as well, although in this case sometimes smaller individuals may have the advantage. Similarly, size is usually directly related to reproductive output, thus putting smaller individuals at a clear disadvantage. Access to different types and amounts of food, refuges, or other resources also often depends on size. Because individuals of different sizes thus make different contributions to population growth, it seems reasonable to expect that populations with different size distributions (e.g., figure 2.5) could well have different dynamics, even if the total biomass of the populations were the same. In the rest of this section we first describe how to quantify the size distributions or “size structures” of populations and then explore the mechanisms that generate size structure, especially population density and competition, and the population-dynamic consequences of size structure. In the section that follows we deal with the rather more complicated problem of modeling size- (or, more generally, stage-) distributed populations.

Not only do individuals typically differ in size within a population, even when all are the same age, but they also often show a particular pattern that is highly skewed, with many small individuals and a few large ones (see figure 2.5). Although plants have probably been the most extensively analyzed with respect to size structure, many other organisms show this pattern also—

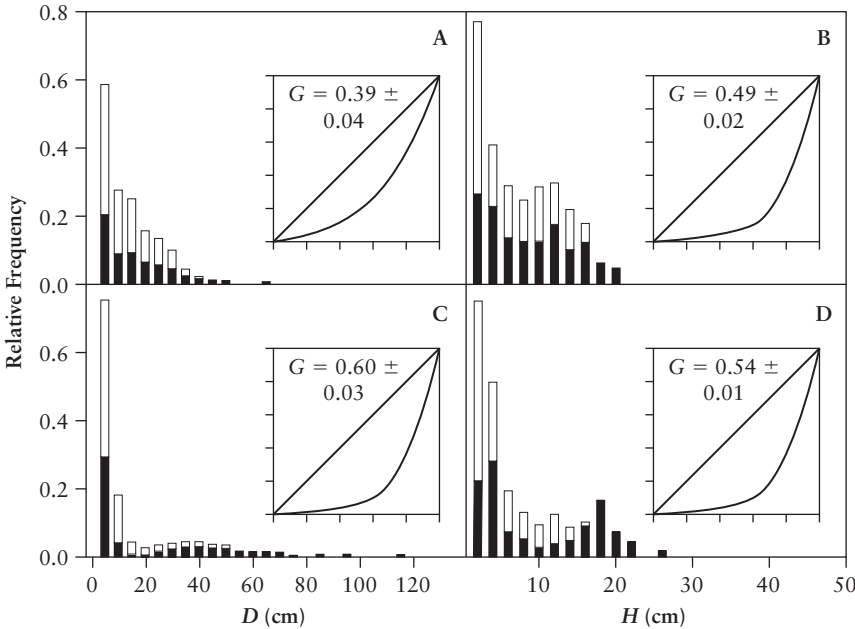


FIGURE 2.5. Size frequency distributions of two taxa of mangrove seedlings in two locations in Mexico. Size is shown as both diameter (D) at breast height (A, C) and height (H) (B, D). G is the Gini coefficient (from Méndez-Alonzo et al. 2012).

for example, fish, corals, wasp colonies, brown algae, and red algae. This particular pattern is often described as a size hierarchy, conveying that the few large individuals have some advantage, such as better access to resources or mates. To quantify size hierarchies, Weiner (1986) borrowed the Gini coefficient from economics, where it is a measure of inequality of distribution. The Gini coefficient is calculated as

$$G = \frac{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{2\bar{x}n(n-1)},$$

where x_i is the size of the i th individual in the population, $\bar{x}$ is the mean of the sizes, and n is the population density.

Where does the broad range of variation in size within so many populations come from? At a most general level, we can write

$$S(t+1) = rS(t), \quad (29)$$

where $S(t)$ is size at time t and r is relative growth rate, now of an individual rather than of a population of individuals as in chapter 1. Note that just as r in a population model is the per capita (per-individual) contribution to population growth, r in an individual growth model is the per-unit size change in size per unit size, that is, the relative growth rate (RGR), and not the absolute (or whole-organism) growth rate. Because this is an exponential form, this equation cannot be true for an extended period of time, but for initial growth of individuals it is frequently a convenient shorthand. Just as we did for population growth, we can restate this equation as

$$S(t+1) = r^t S(0),$$

where $S(0)$ is initial size (e.g., the size of an egg or seed) and t is the time since birth. These variables, along with the relative growth rate, r , therefore collectively determine the size attained by an individual at any time, and all three can be modified by many ecological factors. Although the effects on S_t of the elapsed time since birth or emergence, t , is straightforward, the effects of initial size, $S(0)$, and relative growth rate, r , are more complicated because they are often not independent of each other. For example in plants, larger seeds are often associated with lower initial relative growth rates in the absence of competition (Turnbull et al. 2012), so that these two factors can actually work against each other.

The early literature in plant population ecology often assumed that the existence of size hierarchies was evidence of competition among individuals. However, even populations of individuals in competition with one another will develop a size hierarchy simply because of the exponential nature of early plant growth in biomass (Turner and Rabinowitz 1983), as described by equation 29. If individuals vary even slightly in initial size (or emergence time or relative growth rate), any individual that gets a slight head start will grow at a greater absolute rate, even if at the same per-unit size rate, and the differ-

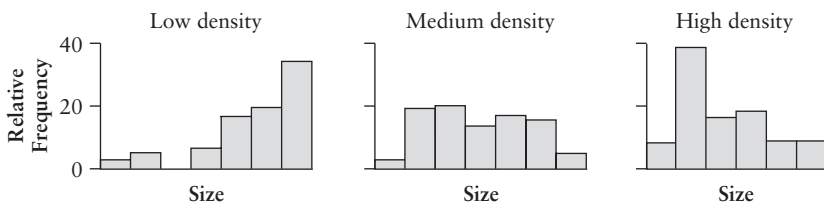


FIGURE 2.6. Intraspecific competition and growth in populations of the limpet *Patella cochlear*. High-density populations have many small individuals and a few large ones; low-density populations have many large individuals and a few small ones (Branch 1975).

ence in size among individuals will continue to increase during the exponential growth phase.

Although competition is not necessarily the cause of size hierarchies, density and therefore the intensity of competition can influence the rate at which hierarchies develop. For example, in a study of the limpet *Patella cochlear*, Branch (1975) found that high-density populations have a large number of small individuals and only a few large ones, whereas low-density populations have a large number of large individuals and only a few small ones (figure 2.6).

The influence of density on size structure takes on three distinct forms. First, individuals that are larger than others may take up a disproportionate share of resources. That is, relative growth rate should increase with size, so that as density increases, the initially larger individuals gain a disproportionate advantage and become much larger, while even slightly smaller individuals do not gain as much. Thus individuals become more and more unequal in their absolute size. This causes the coefficient of variation (or Gini coefficient) to increase with density at a given time or over time at a given density (Weiner and Thomas 1986).

Second, if each individual receives resources in a share proportional to its size, larger individuals gain more resources, but if they are twice as large they obtain only twice as much. Thus all individuals gain a proportional share of resources. This means that relative growth rate is constant with size, and as density increases, every individual obtains proportionately less and grows more slowly. Because the coefficient of variation (or Gini coefficient) increases with size (i.e., the original distribution spreads out), if all individuals grow more slowly, there is less variability at higher densities. This is the opposite of what is expected with disproportionate resource use.

These two forms of resource uptake (proportionate to biomass and disproportionate to biomass) are parallel to what the literature refers to as size-asymmetric and size-symmetric competition in plants and is closely related to the notion of contest and scramble competition, as defined for animals (Hassell 1975; Nicholson 1954). For plants, it has been suggested that the distinction is related to whether the competition is for light or for nutrients. When competition is for light (also called one-sided competition), it is apparent that larger individuals obtain more than their proportionate share of the resource (they

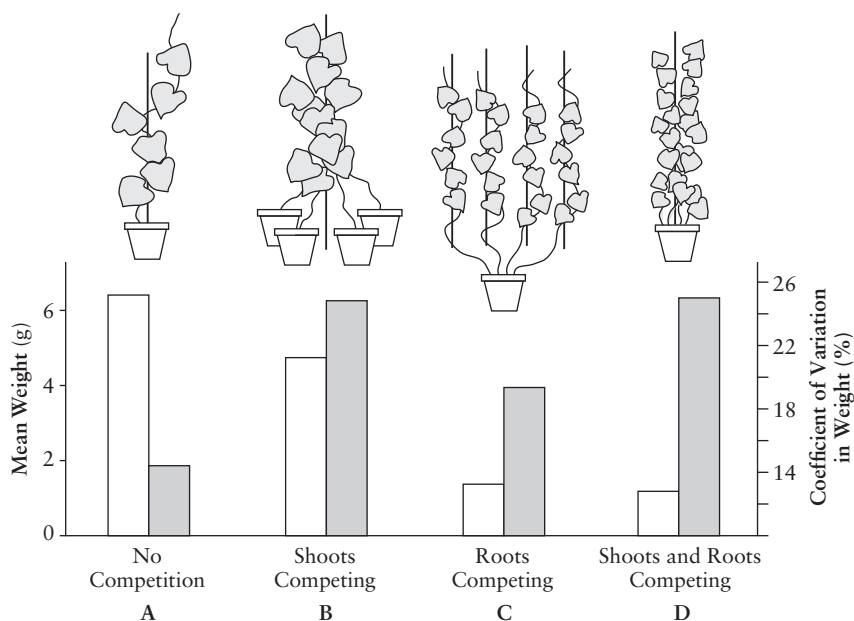


FIGURE 2.7. An experiment on vines (*Ipomoea tricolor*) in which root and shoot competition were separated. Mean mass (open bars) was significantly different ($p < 0.001$) for all comparisons between treatments except C and D. The coefficient of variation (CV) in mass (shaded bars) was significantly different for the comparison of treatments A and B ($p < 0.05$) and of A and D ($p < 0.01$) (Weiner 1986, 1990).

intercept all the light and cast shade on their smaller brethren, who become very little). When competition is for below-ground resources (also called two-sided competition), a bigger individual will have a proportionately larger root system and will more likely consume an amount of nutrients in proportion to its overall size. In a classic experiment, Weiner (1986) used a simple design to separate out effects of one-sided or two-sided competition. His results are illustrated in figure 2.7.

For animals, contest competition is associated with cases of clearly successful and unsuccessful individuals' acquiring resources, such as happens when limited territories or strong social hierarchies exist that determine access to food or mates. Scramble competition occurs when acquisition of resources is either equal or random among individuals. While the winners of contest competition are not necessarily larger individuals, size is often a strong component in determining the outcome, and the consequences in terms of inequitable division of resources and subsequent fitness are clearly the same.

The third way in which competition influences the rate at which hierarchies develop has to do with mortality and self-thinning (see chapter 1). If the smallest individuals are more likely to die (which is almost always true), this will tend to truncate the distribution and thus generate less inequality among surviving plants. So as competition becomes more intense and self-thinning begins to occur, competition results in the reverse of what we expect from

the other two forms of effects of competition on size hierarchies. In plants, as a general rule, competition generates more mortality in the case of competition for light. This is a simple physiological process in that, when a plant is shaded by a competitor such that it is living below its compensation point, it must die. Competition for nutrients generally results in smaller individuals but rarely leads to death.

Consequently, the above can be summarized as a sequence of mechanisms in time for plants. Assume that we start with an even-aged monoculture (e.g., a tree plantation or a field of crop plants). Initially all plants are small and any competition that occurs probably occurs for nutrients and thus is a proportional type. Consequently, higher initial density should lead to less inequality. As the canopy begins closing, competition becomes disproportionate to size and the generation of a size hierarchy is accelerated. Finally, with further canopy closure some plants experience light regimes that are below their compensation points and thus die, reversing the tendency to form a size hierarchy. Obviously this pattern will depend on particular habitats. For example, a particularly low-nutrient situation (e.g., that of a desert) may never proceed to the canopy closure stage.

Regardless of the cause of particular patterns of size structure, the existence of size hierarchies such as that in figure 2.5 seems to be a stabilizing force in population dynamics (Johst 2008; Łomnicki 1988, 2009). Why might this be the case? In simple population dynamics models, we can look at dynamics as output–input graphs or N_{t+1} vs. N_t . In exponential models, this relationship is linear (on a log scale). However, as we discussed in chapter 1, more often than not this relationship is density dependent and thus nonlinear, such that the projection of the population to the next time unit declines at higher densities. Thus, in general we can write

$$N(t + 1) = f(N(t)),$$

where the function f can be very nonlinear with remarkably complicated results, as already shown and as will be discussed in chapter 4. For now suffice it to say that if the function curves downward at high densities, it means that very high-density populations actually produce fewer propagules (eggs, seeds, etc.) than do intermediate-density populations. Depending on how extreme this is, it can lead to population cycles, sometimes very complicated ones (e.g., see figure 2.4).

How does size structure relate to a downward function of population growth rate at high densities? One likely reason for the decline in output at very high input is a lack of size variation. If all individuals are affected equally by the low resources at high density, at a very high density all individuals might be too small to reproduce effectively. On the other hand, if there are size hierarchies that strengthen at high density, there will always be a few large individuals that reproduce adequately and keep total output from declining. Thus it is easy to see that a strong size hierarchy in annual organisms (and perhaps in other organisms as well) tends to stabilize a population over the long run.

Density Dependence in a Stage-Structured Model

More than a decade ago a team of mathematicians and biologists began a collaboration on a series of studies attempting to predict the population behavior of the flour beetle *Tribolium castaneum* (Costantino et al. 1997). Their results have been quite spectacular thus far and warrant introduction here. A fuller presentation of their results will be undertaken in chapter 4 after some more advanced analytical techniques have been introduced. We make the introduction here simply as an example of the way in which (and the complications that arise when) density dependence can be incorporated into a stage-structured model. The analysis of the model necessarily must wait for the development of some additional analytical tools.

Tribolium can be divided into the three stages larva, pupa, and adult. The effect of adults on larvae is complicated because adults both “produce” larvae (the egg stage is not included in this model) and consume larvae. Cannibalism occurs with adults eating both pupae and eggs and larvae eating eggs. Thus we expect a density-dependent effect through this cannibalism in which the production of larvae will be decreased by both more larvae and more adults (because both adults and larvae eat eggs). From the basic biology it is also reasonable to assume that (if the time unit is appropriately chosen) larvae do not remain as larvae from one time unit to the next, nor do pupae, but adults do. Thus the number of larvae at some time point is a function of the number of larvae and adults at the previous point in time, the number of pupae is a function of the number of larvae at the previous point in time, and the number of adults is a function of the number of pupae (some of which will turn into adults, others of which will be eaten by the adults) and the number of adults (the ones that survive). The overall projection model, then, looks like this:

$$\begin{vmatrix} L_{t+1} \\ P_{t+1} \\ A_{t+1} \end{vmatrix} = \begin{vmatrix} 0 & 0 & f_1(L_t A_t) \\ p_{lp} & 0 & 0 \\ 0 & f_2(A_t) & p_{aa} \end{vmatrix} \begin{vmatrix} L_t \\ P_t \\ A_t \end{vmatrix},$$

where L is the number of larvae, P is the number of pupae, and A is the number of adults. The functions f_1 and f_2 stipulate the nonlinear effect of cannibalism on production of larvae by adults and on the survival of pupae to adulthood, respectively. Costantino and coworkers stipulated the functions as

$$f_1 = \frac{b}{e^{c_1 L_t + c_2 A_t}}$$

and

$$f_2 = \frac{b}{e^{c_3 A_t}},$$

where the meanings of the constants c_1 , c_2 , and c_3 are not of particular importance here. Suffice it to say that this simple model defies simple analysis. Its

behavior is remarkably complex, yet with some modern tools of analysis, to be introduced in chapter 4, it is possible to make some sense of the model and, remarkably, the flour beetles seem to behave very much as the model predicts.

APPENDIX: BASIC MATRIX MANIPULATIONS

Matrix Multiplication

A matrix is a table of numbers, and the mathematical manipulation of matrices derives from analysis of systems of linear equations. Thus, for example, if we have the system

$$Y_1 = a_1X_1 + b_1X_2 + c_1X_3, \quad (\text{A1a})$$

$$Y_2 = a_2X_1 + b_2X_2 + c_2X_3, \quad (\text{A1b})$$

and

$$Y_3 = a_3X_1 + b_3X_2 + c_3X_3, \quad (\text{A1c})$$

simply as a matter of convenience we can group the X_i s to the right and more easily visualize the structure of the system as follows:

$$\begin{vmatrix} Y_1 \\ Y_2 \\ Y_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \begin{vmatrix} X_1 \\ X_2 \\ X_3 \end{vmatrix}.$$

Here we have three matrices: first a matrix with a single column (the Y s), second a square matrix with three columns (sometimes referred to as the detached coefficient matrix), and third a matrix with a single column (the X s). Sometimes the various matrices are simply referred to using a single letter, but it is customary when speaking of matrices to put them in boldface type, so the above equation could be

$$\mathbf{Y} = \mathbf{AX}. \quad (\text{A2})$$

Equation A2 is identical to the system A1, except it is obviously more compact. This way of writing a set of linear equations becomes especially convenient when dealing with large systems. A system of 500 equations could be just as easily written as equation A2.

Equation A2 is interpreted in the same way as any other equation—the matrix Y (when there is only a single column, we sometimes refer to the matrix as a “column vector”) is equal to the column vector X multiplied by the matrix A . But multiplying matrices is not as simple as multiplying scalars (nonmatrix variables—that is, regular numbers). We basically need to perform an operation such that we obtain the system A1 from the equation A2. The basic rule is to multiply each element in the first row of the square matrix by the corresponding element in the column vector and sum the results (i.e.,

$a_1X_1 + b_1X_2 + c_1X_3$). That gives us the first element in the Y vector (i.e., $Y_1 = a_1X_1 + b_1X_2 + c_1X_3$). Next we multiply each element in the second row of the square matrix by the corresponding element in the column vector and sum the results (i.e., $a_2X_1 + b_2X_2 + c_2X_3$). That gives us the second element in the Y vector (i.e., $Y_2 = a_2X_1 + b_2X_2 + c_2X_3$). We follow the same procedure to get the third element in the Y vector. Such is the relatively simple process of multiplying a column vector by a square matrix.

Frequently it is necessary to multiply square matrices by square matrices. The process is really nothing more than multiplying the square matrix by successive column vectors. Take the simple example of a pair of two-by-two square matrices,

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \text{ and } \begin{vmatrix} c_1 & d_1 \\ c_2 & d_2 \end{vmatrix}.$$

Begin the process of multiplication by simply considering the first column of the second matrix (i.e., the column with the c s) being multiplied by the square matrix to its left. Following the procedure outlined above, we have $a_1c_1 + b_1c_2$ as the first element in the first column of the resultant matrix and $a_2c_1 + b_2c_2$ as the second element in the first column of the resultant matrix, or

$$\begin{vmatrix} (a_1c_1 + b_1c_2) & \text{---} \\ (a_2c_1 + b_2c_2) & \text{---} \end{vmatrix},$$

where the horizontal lines indicate the calculations that are still to be made. Now repeat the process for the second column in the second matrix (the one with the d s) so that you get $a_1d_1 + b_1d_2$ for the first element in the second column of the resultant matrix and $a_2d_1 + b_2d_2$ for the second element in the second column of the resultant matrix, so that the final result is

$$\begin{vmatrix} (a_1c_1 + b_1c_2) & (a_1d_1 + b_1d_2) \\ (a_2c_1 + b_2c_2) & (a_2d_1 + b_2d_2) \end{vmatrix}.$$

A moment's reflection should convince you that premultiplication is not equivalent to postmultiplication (that is, in multiplying the two matrices A and B together you must specify whether you are multiplying AB or BA —the results will not generally be the same). Furthermore, it is possible to multiply matrices together only if the number of columns of the first matrix is equal to the number of rows of the second.

Matrix Addition and Subtraction

Adding and subtracting matrices is a far simpler affair than multiplying them. Just add or subtract the equivalent elements in each matrix. Note that you can add or subtract matrices only if they have the same number of rows and columns. The following example should make the process crystal clear. Subtract matrix A from matrix B below.

$$A = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \text{ and}$$

$$B = \begin{vmatrix} c_1 & d_1 \\ c_2 & d_2 \end{vmatrix}, \text{ so}$$

$$A - B = \begin{vmatrix} (a_1 - c_1) & (b_1 - d_1) \\ (a_2 - c_2) & (b_2 - d_2) \end{vmatrix}.$$

The Identity Matrix

Recall from basic arithmetic the identity element. We need an element such that when we multiply it by any other element, the result is that same element again. That is,

$$aI = a,$$

where I is the identity element. In elementary arithmetic we all recognize $I = 1$. Any number multiplied by 1 gives us that element back again. A similar requirement exists in matrix algebra. That is, we need an element such that $AI = A$ (where A and I are matrices, as indicated by their boldfaced status). Recalling the basic rules of matrix multiplication, it does not take much thought to convince yourself that a matrix with 1s on the principal diagonal and 0s everywhere else will act as the identity element for any square matrix (with the same number of rows as the identity matrix). That is, consider the three-by-three matrix

$$A = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

If we perform the multiplication of AI , we obtain

$$AI = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

The Determinant of a Matrix

The determinant of a matrix is an important concept, the significance of which is difficult to convey in any sort of intuitive fashion. The concept derives from solving systems of linear equations. Consider, for example, the following set of equations:

$$k_1 = a_1X_1 + b_1X_2 \tag{A3a}$$

and

$$k_2 = a_2X_1 + b_2X_2. \quad (\text{A3b})$$

Solve the first equation for X_1 , as follows:

$$X_1 = (k_1/a_1) - (b_1/a_1)X_2.$$

Now substitute the value of X_1 (i.e., the right-hand side of the above equation) into A3b and solve for X_2 , obtaining

$$k_2 = a_2[(k_1/a_1) - (b_1/a_1)X_2] + b_2X_2,$$

and rearrange it so as to put X_2 on the left-hand side:

$$X_2 = \frac{a_1k_2 - a_2k_1}{a_1b_2 - a_2b_1}, \quad (\text{A4})$$

Look closely at the denominator. If we arrange equations A3a and A3b as a matrix equation, we get

$$\begin{vmatrix} k_1 \\ k_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \begin{vmatrix} X_1 \\ X_2 \end{vmatrix},$$

where the square matrix of constants is referred to as the detached coefficient matrix.

If we now multiply the elements of the principal diagonal together (i.e., a_1b_2) and subtract from that the product of the two off-diagonal elements (a_2b_1), we have precisely the denominator of equation A4. This value is referred to as the determinant of the matrix, and symbolized as $\text{Det}A$, for the determinant of the matrix A . So we have

$$\text{Det } A = a_1b_2 - a_2b_1,$$

where the matrix A is

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

Look at the numerator of A4. Now form the matrix

$$\begin{vmatrix} a_1 & k_1 \\ a_2 & k_2 \end{vmatrix},$$

which is to say, replace the second column of the detached coefficient matrix with the column of constants k_1 and k_2 . Recall the basic definition of the determinant—multiply the elements of the principal diagonal, and subtract from them the product of the off-diagonal elements to obtain

$$\text{Det} \begin{vmatrix} a_1 & k_1 \\ a_2 & k_2 \end{vmatrix} = a_1k_2 - a_2k_1,$$

which you will recognize as the numerator of equation A4. So we could generalize and say that if we form the matrix A_i , in which the i th column (in this case $i = \text{either } 1 \text{ or } 2$) of the detached coefficient matrix has been replaced

with the column of constants (i.e., the elements of the column vector on the left-hand side of the equation), we have a general equation,

$$X_i = \text{Det } A_i / \text{Det } A.$$

This equation is known as Cramer's rule (Cramer was a Frenchman, so it is Kra-MAY's rule). The amazing thing about it is that it applies to any system of linear equations. If you just solve an arbitrary system for one of the variables (say the fourth variable), you come up with a ratio that is the ratio of two determinants, the numerator being from the matrix that had its fourth column substituted with the column of constants and the denominator being the determinant of the detached coefficient matrix itself. Of course calculating the determinant for a matrix larger than two by two is more complicated but is not really required of the reader of this book.

3

Applications of Simple Population Models

In the previous two chapters we showed how simple mathematical models can illustrate general principles about population dynamics. In this chapter we illustrate the application of some of these models and associated tools to address a number of different kinds of problems. We start with the basic models of population dynamics from chapter 1 and apply them to the problem of the evolution of life histories and then move to using the structured models of chapter 2 to describe more complex patterns in life histories. We then turn to applications of population projection matrices both for natural resource management and for conservation.

The most obvious uses of population-dynamic models are to examine the range of kinds of dynamics we might observe in real populations and, more directly, to make at least short-term population projections. However, we start this chapter by focusing on a very different and important use of population-dynamic models, the study of life history evolution. This activity has a long history in ecology, from the classic studies of Lack (e.g., 1947) and Cole (1954) to the present time. Cole (1954), for example, showed that adding one offspring in a previous year could have much greater effects on long-term population growth than adding several more offspring in a current year, which suggests strong selection for earlier reproduction. Or, if evolutionary pressure were exerted on individuals within a population to be more efficient in their use of resources, modeling that population with the logistic equation would begin with the assumption that evolutionary pressure was causing the carrying capacity, K , to increase. On the other hand, if evolutionary pressure on the individuals in a population caused each individual to require more space in the environment, we might model that situation with a logistic equation again, but this time incorporating a decreasing K . In either case we can ask what will happen to the population over the long run if these evolutionary pressures are brought to bear. This is the central problem of life his-

tory evolution—how life history traits evolve—and it is usually approached through the vehicle of asking how proposed patterns of selection pressure (which inevitably act on individuals) will affect the population as a whole. These results are then interpreted using the per capita long-term population growth rate as a measure of individual fitness.

Life History Analysis

Life history traits are traits that affect the basic survival and reproductive schedules of organisms (e.g., age-specific birth rate, stage-specific growth rate, crude death rate, size at birth, number of offspring, size of offspring, longevity, etc.). Many of these traits come directly out of the population projection matrices introduced in chapter 2 to study population dynamics. There we saw that changing the transition probabilities in a matrix can change the population growth rate, and we used tools such as sensitivity and elasticity to analyze those effects. The important point for this section is that the variables we used to quantify the population growth rate were all used on a per-individual (per capita) basis. That means that the full mathematical machinery and tools presented in the previous chapters also can be used to analyze the effects of changes in life history on individual fitness. And because natural selection acts mainly on individuals, we can thus explore the potential evolutionary consequences of different life history traits.

The key assumption that underlies life history theory is that there are trade-offs among these life history components. If such trade-offs did not occur, it is fairly obvious what natural selection would ultimately produce—immortal individuals would start reproducing at birth and continue producing many large offspring in perpetuity. This is obviously not what happens, and the reason is that there are costs associated with reproduction. For example, a high investment in early reproduction early in life likely reduces an individual's subsequent reproductive success, or producing large offspring means that fewer of these offspring can be produced.

In this section we concentrate on what could be regarded as the main trade-off—the relationship between investment in survivorship versus investment in reproduction. First we describe the classic life history ideas of r versus K selection (selection for intrinsic rate of natural increase versus carrying capacity), which come directly from the logistic equation. Although these ideas imply a trade-off between survivorship and fecundity, this is not explicit, because birth and death rates are not parameters in this equation. Second we ask how to quantify the cost of reproduction that is the ultimate basis of this trade-off: a high rate of current fecundity should mean a lower rate of survival (due to less allocation of limiting resources to maintenance, growth, and defense) and/or subsequent reproduction. The assumption of a cost of reproduction is the most fundamental assumption of life history theory, but testing it is not so simple. Third, assuming that there is a cost to reproduction, we ask what is the optimal schedule of reproduction (when and how much each time) and, consequently, survival.

Investment in Survivorship versus Reproduction: The r - K Continuum

Life history evolution can be traced to the early work of Cole and Lack (see above), but its birth in modern form can be seen in the musings of Cody (1966) and MacArthur and Wilson (1967). The basic formulation follows the lead of the logistic equation. In that equation (see chapter 1) there are two parameters, r and K , the intrinsic rate of natural increase and the carrying capacity, respectively. MacArthur and Wilson (1967) further noted that at low density, the per capita population growth rate is primarily influenced by the value of r , the intrinsic rate of natural increase. (When N is small, $[K - N]/K$ is almost equal to $K/K = 1$, so dN/Ndt approximates r .) However, at high density, the per capita population growth rate is primarily influenced by the value of K , the carrying capacity. Because the per capita population growth rate is a reflection of individual fitness, this suggests that natural selection will act primarily on traits that increase r at low density and on traits that influence K at high density. Hence the appellation r - K selection.

Although the formal basis of r - K selection theory is strictly that of density-dependent selection (Boyce 1984), the idea rapidly broadened to consider any environmental conditions that might lead to low versus high density. Thus r selection was assumed to be the dominant mode of selection early in succession, in highly disturbed environments, and in unpredictable environments, while K selection was assumed to be the dominant mode of selection in more stable and predictable conditions, in which density is assumed to be high (table 3.1). r and K selection then became associated with the tendency of organisms to be at one extreme or another in their pattern of evolutionary investment in reproduction and survival (Cody 1966; Pianka 1970). Some species tend to produce large numbers of offspring, begin reproduction early, and die young; these are traits associated with r selection. Other species tend to produce small numbers of offspring but invest considerably in parental care, begin reproduction later in life, and live relatively long lives, associated with K selection. Pianka (1970) suggested that insects in general are at one end of the spectrum while mammals are at the other. The general pattern of attributes associated with each end of this continuum are listed in table 3.1. The idea in its most general form is that we can imagine populations existing on an environmental continuum ranging from unpredictable/harsh to constant/mild. Although such terminology is vague enough to encourage some to fit any observation whatsoever into the scheme, it is also vague enough to cause others to criticize its utility in the first place (e.g., Hairston et al. 1970; Wilbur et al. 1974). Nevertheless, it is at least an important historical milestone in the field of life history evolution.

Although the first of these sets of traits is clearly associated with high values of the intrinsic rate of natural increase, a major critique of the r/K paradigm is that the opposite set of traits has little to do directly with the population carrying capacity, K . Instead, K should be most directly related to individual traits of resource use efficiency that would enable more individuals to persist in an environment with a given amount of resources. That is, K is not a demographic parameter, so it doesn't fit easily into a theory of evolution

TABLE 3.1. Some Attributes of r - and K -Selected Species and the Environmental Characteristics That Select for Them (Pianka 1970)

Attribute	r selection	K selection
Mortality	Variable and unpredictable	More constant and predictable
Population size	Variable, below carrying capacity	Constant, close to carrying capacity
Intra- and interspecific competition	Variable, often weak	Usually strong
Selection favors	Rapid development Early reproduction Small body size Semelparity	Slow development Delayed reproduction Large body size Iteroparity
Length of life	Usually shorter	Usually longer
Leads to	High productivity	High efficiency

of life history traits. This leads to the general criticism of the entire paradigm of r and K selection that the characteristics upon which selection acts are not really related to these particular phenomenological categories (i.e., intrinsic rates of natural increase and carrying capacity) but are rather characteristics associated with birth rates and death rates (Hairston et al. 1970; Wilbur et al. 1974). Nevertheless, the r and K selection paradigm remains a staple of introductory ecology textbooks.

Going beyond the r/K ideas, we inevitably must deal with models that explicitly incorporate birth and death rates and are often based on structured populations of some kind. These can be considerably more complicated than the simple logistic. The general strategy is to focus on birth and death schedules and assume that natural selection maximizes the per capita growth rate of the population (Metcalf and Pavard 2006). The main conceptual framework that has been used is that of trade-offs. We begin by assuming that no organism can do everything. On the one hand, there are phylogenetic constraints on natural selection (which is why there are no six-legged vertebrates or four-legged insects). More immediately, there are costs of reproduction and survivorship such that perhaps high investment in reproduction early in life reduces subsequent reproductive success or making large offspring means that fewer can be made. So we have to ask what are the optimal reproductive and survivorship schedules given the relevant costs and how those optima change as the costs change. First we ask, Is there a cost of reproduction, and how should it be measured? Second, assuming there is such a cost, how do reproductive schedules evolve, and are they optimal?

The Cost of Reproduction

A great deal of attention has been given to this topic in the past, for both plants (Ashman 1994) and animals (Roff 1992; Stearns 1992). We can divide the approaches taken into studying the costs of reproduction as, first, phenotypic or genetic, and within each of these as correlational or experimental. Phenotypic approaches give only an idea of intrinsic physiological constraints. If there is no genetic basis for these constraints, it is not a topic for analysis using the theory of natural selection. However, if there are strong physiological constraints, it seems a reasonable assumption that the constraints are pleiotropically related (i.e., that the same gene complexes are responsible for increased reproduction and decreased maintenance). Many studies show a negative correlation between fecundity in year x with survival in year $x + 1$, as would be expected by a trade-off model. However, many other studies do not show this pattern, and indeed some studies show the opposite. This variability underscores the problems inherent in trying to elucidate deeper meanings from simple correlations. Frequently it seems to be the case, at least with sessile organisms, that an individual in a bad microsite has both low fecundity and low survivorship, while an individual in a good microsite has both high fecundity and high survivorship. Consequently, an empirical correlation actually shows a positive relationship between fecundity and survivorship.

Many studies have experimentally manipulated phenotypes, and in these, too, the results have been variable (Ashman 1994; Golet et al. 2004). The underlying theory strongly suggests that there should be substantial cost to reproduction, yet frequently in experimental situations such an expectation is not met. Ashman argues that part of the problem is the difficulty in measuring investment in current reproduction. It is not actually fecundity that determines the cost but the total investment in reproduction (e.g., the cost of attracting mates through behaviors or structures such as flowers). Also complicating investigations is the need to use a currency that is limiting to the organism; although biomass is commonly used to compare allocation to different functions as a proxy for energy, the limiting resource may be something quite different, such as a particular nutrient or even time.

Studies of genetic correlations and artificial selection experiments have become increasingly common over the past decade and are undoubtedly the best way of evaluating costs of reproduction in the context of deducing potential for evolutionary change. Perhaps the classic case is that of Rose and Charlesworth's (1981) natural selection experiments with *Drosophila melanogaster*, in which two lines were selected for. In line O, the flies were not allowed to successfully breed until later in life, while in line B flies were not allowed to successfully breed later in life. The expectation, then, would be that O flies would evolve to be late breeders and B flies early breeders. Not only did O flies turn out to be later breeders than B flies; they also evolved the capacity to live longer (and therefore demonstrated a genetic correlation between length of life and time of breeding). The "cost" in this case of early reproduction was shortened life span (see figure 3.1).

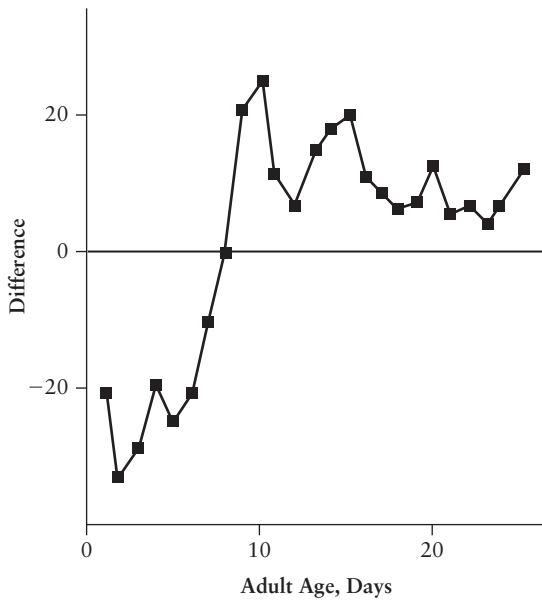


FIGURE 3.1. Difference in per capita egg production between the O lines and B lines from Rose and Charlesworth's (1981) experiment.

In addition to their importance in evolutionary theory, questions about the cost of reproduction may have important practical consequences. For example, Jackson (1980) has long promoted the idea that it would be useful for natural systems agriculture to breed grasses that are both perennial and high yielding (because natural prairies are based on perennial grasses, this is a clear goal resulting from the philosophy of natural systems agriculture). But the basic evolutionary theory of trade-offs between reproduction and maintenance would suggest that this might be an impossible goal. All domesticated grasses thus far are annuals, and one might think that the cost of maintaining rootstocks permanently would mitigate against allocating more energy to the seeds (which is what we want to harvest). However, Jackson and Dewald (1994) demonstrated that the perennial grass *Tripsacum dactyloides* appears to compartmentalize energy allocation such that significant changes in allocation to reproduction can occur with no changes in allocation to maintenance activity (the compartmentalization appears to occur because of the autonomous development of reproductive tillers). This may suggest that the assumptions of a simple trade-off model are too simplistic (Lande 1982).

Optimal Reproductive Schedules

The fundamental question in life history evolution is when and how much an organism can reproduce. This may include, for example, age at first reproduction, frequency of reproduction, allocation to each bout of reproduction, and how these are affected by environment—overall favorability, fluctuations, density, and so on. It is commonplace in studies of these questions to use phenotypic optimality models. These models assume that sufficient genetic variability exists and is structured in such a way that there are no constraints on

reaching optima. This is obviously not true, and much more detailed models and analyses exist that do incorporate details of genetic structure. In fact, much of contemporary research on life history evolution is focused on the integration of the genetic and physiological mechanisms underlying life history traits (Flatt and Heyland 2012). The role of phylogeny in constraining evolutionary trajectories has also become central to much of modern life history analysis (e.g., Bielby et al. 2007; Burns et al. 2010; Silvertown et al. 1997; Warne and Charnov 2008).

Nevertheless, we focus here on very simple optimality models to illustrate some general principles and approaches. The point of phenotypic optimality models, especially at the very coarse scale intended here, is to predict general patterns across species and environments that we expect to be roughly true and thus suggest what specifically empiricists should be measuring that will help explain these ecological patterns. Bull and Wang (2010) also argue that phenotypic optimality models can be very useful complements to experimental evolution in exploring the genetics of adaptation.

Assuming that, despite the difficulty in measuring it, there is a cost to reproduction, what is the optimal reproductive schedule? It is clear from basic biology that a plant or animal devotes some accumulated energy to reproductive activities and the rest to maintenance and growth activities. How much energy should be allocated to reproduction versus maintenance and growth during a given season? The most elementary formulation to answer this question has no structure associated with it and involves the simple exponential map (see chapter 1), namely

$$N_{t+1} = \lambda N_t,$$

where N is the density or biomass of the population. The parameter λ can be broken into two parts,

$$\lambda = b + p,$$

where b is per capita births and p is probability of survival (note that this is the reverse of the death rate; $1 - p$ is the probability of death). One can conceive of an organism as having to “decide” (short for natural selection acting on the problem) how much energy to devote to b , reproduction, versus other functions that will enhance p . This means that both b and p must be cast as functions of something else. Following the formulation of Schaffer and Gadgil (1975), we allow both b and p to be functions of energy devoted to reproduction. So we write

$$\lambda = b(E) + p(E),$$

where E is the fraction of total energy devoted to reproduction (so E energy goes to b , and $1 - E$ energy goes to p). How might we expect b and p to vary with E ? In general, b increases and p decreases with E (at the extreme, if $E = 100\%$, p has to be 0; there is nothing left for maintenance because it has all been used up in reproduction). The expectation is that natural selection will act to maximize the value of λ , that is, the sum of b plus p . Using this model

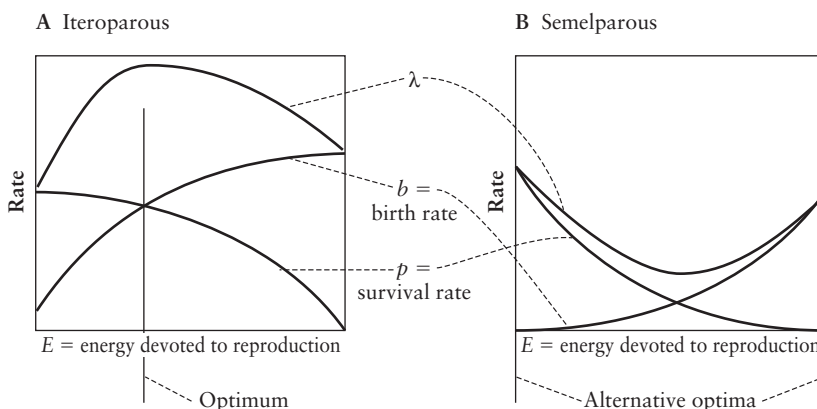


FIGURE 3.2. Alternative solutions for the Schaffer model of energy allocation. b is per capita birth rate, p is per capita survival rate, and λ is their sum, per capita population growth rate.

Schaffer (1974) discusses two general situations: either both functions are concave (i.e., have a negative second derivative, as in figure 3.2A) or both are convex (i.e., have a positive second derivative, as in figure 3.2B).

The result of this simple thought experiment is that we expect there to be two qualitatively distinct reproductive strategies in nature, one in which there are two alternatives—(1) no reproduction or (2) the organism puts everything into reproduction (the semelparous strategy, shown in figure 3.2B)—and one in which there is a single optimum—the organism puts some energy into reproduction and some into maintenance every year (the iteroparous strategy, shown in figure 3.2A). This, of course, corresponds to the situation we already know occurs in the real world, but here we have a model that suggests something of an underlying mechanism that gives rise to the strategy. If both curves are concave, there is a maximum at an intermediate E , which is the iteroparous solution. When both are convex, the maximum is at either 0 or 100%, which is the semelparous solution.

Given these two dramatically different solutions that emerge from the underlying shape of the functions, it is of interest to ask what ecological factors give rise to either concave or convex functions. The arguments are usually focused on the fecundity curve, because survival can be thought of as everything but fecundity. A concave fecundity function implies diminishing returns for the energy devoted to reproduction. For example, if competition for pollinators decreases the proportion of flowers pollinated as more flowers are produced, a concave fecundity function will result. This means that although the total *number* of pollinated flowers will continue to increase with increasing energy devoted to reproduction, the *probability* that any given flower will be pollinated declines. Or, to take an animal example, if individual birthweight decreases when more individuals are produced (as it must in many animals, such as mammals), such that spontaneous abortions occur with large litter expectancies, a concave fecundity will again result.

On the other hand, a convex function will result from an accelerated accumulation of births with increasing energy devoted to reproduction. For instance, if the production of large numbers of flowers on a single plant attracts disproportionately more pollinators than a smaller number of flowers, reproduction will increase nonlinearly with more energy devoted to reproduction (i.e., producing more flowers). In this case, not only will more flowers be produced with more energy allocated to reproduction, but each flower will also have a higher probability of being pollinated. Another example might arise in a situation in which seed predators are important in the system. With low numbers of seeds produced, not many seedlings are produced because almost all are eaten by seed predators. However, when seed production increases, there is a tendency for the seed predators to become satiated, thus creating the critical nonlinearity (Silvertown 1980).

As a general rule, we can say that some sort of facilitation is involved when there is a convex function, whereas some sort of intraspecific competition is involved when the function is concave. An interesting switch between the two has been found in the scarlet gilia (*Ipomopsis aggregata*) (Paige and Whitham 1987). At low elevations this species is a classic semelparous species, but at higher elevations it sometimes can be iteroparous. The presumed reason for this switch is the lower abundance of pollinators, which implies more competition for pollinators, a force that would tend to generate a convex curve. Paige and Whitham (1987) garnered further evidence of this mechanism by showing that more iteroparity was observed when they excluded pollinators from plants. However, they also found a greater tendency for iteroparity when they removed flower buds early in the season, a treatment that should have reduced the competition for pollinators.

If it is the case that we are dealing with an iteroparous situation, finding the optimal energy level to allocate for reproduction is simple. Given the basic equation for growth rate,

$$\lambda = b(E) + p(E),$$

we wish to find the maximum value of λ with respect to E . Thus we compute

$$\frac{d\lambda}{dE} = \frac{db}{dE} + \frac{dp}{dE}$$

and set

$$\frac{d\lambda}{dE} = 0$$

(in addition, the second derivative must be negative to ensure that it is a maximum rather than a minimum, which is the case given our assumption that the situation is iteroparous; see figure 3.2) to obtain

$$\frac{db}{dE} = -\frac{dp}{dE}. \quad (1)$$

We thus conclude that when the rate of change of birth rate with respect to reproductive energy is equal to the (negative) rate of change of survivorship

with respect to reproductive energy, the population growth rate is optimized (i.e., where the two curves cross in figure 3.2).

The basic model can be modified in various ways to answer particular questions. For example, what might be the effect of environmental variability acting on the birth rate? What might we expect for the optimal strategy? Suppose we have two environments that occur in a coarse-grain fashion (that is, a given year is either a good year or a bad year). The birth rate can be thought of as being multiplied by an environmental factor such that $b(1 + s)$ is the birth rate in a good year and $b(1 - s)$ is the birth rate in a bad year. Then we have

$$\lambda_g = b(1 + s) + p$$

and

$$\lambda_b = b(1 - s) + p,$$

where the g subscript refers to good years and the b subscript refers to bad years. If we presume that the good years and the bad years occur with the same frequency, we can express the overall population growth as

$$\lambda^2 = \lambda_g \lambda_b = [b(1 + s) + p][b(1 - s) + p] = b^2 + 2bp + p^2 - b^2 s^2,$$

and the derivative of population growth rate with respect to the energy allocated to birth becomes

$$\frac{d(\lambda^2)}{dE} = 2b \frac{db}{dE} + 2p \frac{db}{dE} + 2b \frac{dp}{dE} + 2p \frac{dp}{dE} - 2bs^2 \frac{db}{dE}.$$

To find the optimum energy allocation, we again set this derivative equal to zero and rearrange to obtain

$$\frac{db}{dE} \left(1 - \frac{bs^2}{b + p} \right) = - \frac{dp}{dE}.$$

And because

$$\frac{bs^2}{b + p} > 1,$$

by definition (assuming that the birth rate is positive), we compare equation 2 with equation 1 and conclude that the energy devoted to reproduction should decrease when variable environments affect birth rates (i.e., when $s > 0$). It pays to reproduce at a lower rate if doing so increases the chance of surviving to reproduce again, a phenomenon that has come to be known as bet hedging. The notion of bet hedging has become a major theme in the analysis of life histories (Childs et al. 2010).

An important question arises when dealing with this sort of theory: what, precisely, should one measure in nature to test the theory? The theory says quite unambiguously that one should measure b and p as a function of E , the proportion of total energy devoted to reproduction. However, because one

can never know all of what a plant or animal is doing physiologically, it is difficult to be certain that one is measuring all the components of reproduction. Often the best that can be done in field studies is to use surrogates for the components of the model. For example, Schaffer and Schaffer (1977) used the number of seeds per pod (i.e., the probability of being pollinated, an index of fecundity, which would most likely be correlated with λ) as a function of inflorescence size (a surrogate for energy spent in reproduction) for several iteroparous and semelparous species. The model predicts that there should be a relationship between these two variables for semelparous species but not for iteroparous species. Schaffer found that the relationship between seeds per pod and inflorescence size was indeed correlated significantly for semelparous species but not for iteroparous ones, precisely as predicted by the model.

A similar analysis can shed light on when a plant is expected to be an annual versus a perennial (Schaffer and Gadgil 1975). For an annual plant we can write

$$N_{t+1} = c(bN_t),$$

where b is the number of offspring per individual, as before, and thus bN_t is the total number of juveniles produced, and c is the juvenile survival probability. A perennial plant carries over some of the adults from year to year and thus can be represented as

$$N_{t+1} = c(bN_t) + pN_t,$$

where p is the proportion of adults surviving into the next time period. We thus see that the difference between the two strategies is simply setting the adult survival rate to zero (the definition of annual). So for the annual plant we can write

$$\lambda_a = cb_a$$

and for the perennial

$$\lambda_p = cb_p + p,$$

and using the previous logic that natural selection maximizes λ , we conclude that a plant should evolve perenniality (that is, the condition for $\lambda_p > \lambda_a$) when

$$cb_p + p > cb_a,$$

which can be rewritten as

$$b_a > b_p + p/c.$$

This forces us to conclude that the fecundity of an annual has to be greater than the fecundity of a perennial for the annual to be the favored strategy. Furthermore, if p is approximately equal to c we have

$$b_a > b_p + 1,$$

which says that all an annual has to do is produce one additional seed than a perennial and it will have greater fitness. Why, then, are there any perennials?

Applications of Population Projection Matrices

The main problem associated with applying structured models in nature is one of estimating the values of the parameters. Unfortunately it is very difficult to treat this subject in a general fashion because each organism has special features that need to be taken into account in estimating transition probabilities and fecundities. However, several points need to be noted.

The basic element of structured population models is the probability of transition from one stage to another, which is ideally obtained by marking and following individuals over time. For example, if there are $N_i(t)$ individuals in stage i at time t and $N_{i+1}(t+1)$ individuals in stage $i+1$ at time $t+1$, we *cannot* conclude that the transition probability from i to $i+1$ is the simple ratio $N_{i+1}(t+1)/N_i(t)$ unless the stage is strictly age (i.e., it is a Leslie–Lewis model). However, if we have $N_i(t)$ individuals marked (all in stage i at time t) so as to recognize those particular individuals in the future and we find that of those $N_i(t)$ marked individuals, $N_{i+1}(t+1)$ of them show up in the next time interval in the $i+1$ th stage, it does indeed seem that we may conclude that $P_{i,i+1}$ (the transition probability from stage i to stage $i+1$) is equal to the simple ratio $N_{i+1}(t+1)/N_i(t)$. Such a simple calculation is in fact not necessarily completely correct because the calculation depends on the assumption that the population in the stage category i to $i+1$ is distributed according to a stable stage distribution (Vandermeer 1975). This assumption is most likely to be met if the definition of *stage* is as narrow as possible. However, making stage definitions narrow limits the amount of data available to make the estimate. This basic contradiction is inherent in any study attempting to instantiate a structured population model.

Estimation of transition probabilities and fecundities from following individuals over time results in what is frequently referred to as a dynamic life table (or a cohort life table). An alternative approach, the static life table, uses a snapshot at one point in time of how many individuals are in each stage or age category and uses these numbers to estimate survivorship and reproduction.

The Dall's Mountain Sheep: A Static Life Table

Consider, for example, the classic case of the Dall mountain sheep (*Ovis dalli*), originally analyzed by Murie (1944) (see Deevey 1947 for a summary), with the results reproduced below as table 3.2. The number of individuals in each age category was estimated by the ages of 608 carcasses encountered in Mount McKinley National Park (now Denali National Park). If we assume that their survival and number of births were constant year to year, we can calculate survival rates from these data. This can be true only if the total number of births in the population was the same, that is, if λ was near 1 and the survival rates were constant year to year. Such assumptions are rarely met in natural populations, but frequently they are made so as to get at least some information about a critical population.

Such a life table is constructed by summing up all the death records and supposing that the total number of deaths was the number of individuals in

TABLE 3.2. Static Life Table Based on 608 Carcasses of Dall’s Mountain Sheep Encountered by Murie (1944) in Mount McKinley National Park

Age interval, years	Number dying during age interval	Number surviving at beginning of age interval	Number surviving as a fraction of original cohort
0–1	121	608	1
1–2	7	487	0.801
2–3	8	480	0.789
3–4	7	472	0.776
4–5	18	465	0.764
5–6	28	447	0.734
6–7	29	419	0.688
7–8	42	390	0.640
8–9	80	348	0.574
9–10	114	268	0.439
10–11	95	154	0.256
11–12	55	59	0.096
12–13	2	4	0.006
13–14	2	2	0.003
14–15	0	0	0

the population at the beginning (thus the sum of numbers in column 2 is 608, which was the number surviving at the beginning of the first age interval). Then each category’s death record was subtracted to produce the estimate of the number surviving at the beginning of the age interval. Finally, each number in column 3 was divided by the original number in the cohort to produce the number surviving as a fraction of the original cohort. This is the classic way of elaborating a static life table and can be used to estimate survivorship probabilities from stage to stage.

Such static life tables are of limited use because their construction is dependent on the assumption that the population is at a “stationary” age distribution, which is to say that the eigenvalue is 1.0 and the population is at a stable age or stage distribution. Dynamic or cohort life tables are generally more interesting in that one can ask whether the population is growing or declining, whether the stable age distribution has been reached, what the elasticities are, and a host of other questions.

Palo de Mayo: A Dynamic Life Table

Consider the growth of a population of *Vochysia ferruginea*, a lowland tropical rain forest tree. Boucher and Mallona (1997) report on the growth of a population of this species subsequent to a very large storm. The storm was a hurricane that devastated the entire forest of southeastern Nicaragua in 1988, so for the proximate five years it might be reasonably expected that the population would experience its “maximum” possible growth rate, that is, be relatively free of density-dependent effects. Classifying the population into five

stages (seedling, small sapling, large sapling, young adult, and adult), marking individuals, and following the same individuals over a five-year period, the researchers constructed the following projection matrix:

Stage	Seedling	Small sapling	Large sapling	Small adult	Large adult
Seedling	0.209	0	0	35.60	70.10
Small sapling	0.010	0.653	0.020	0	0
Large sapling	0	0.170	0.407	0	0
Small adult	0	0	0.570	0.731	0
Large adult	0	0	0	0.266	0.997

From this projection matrix they calculated that $\lambda = 1.156$. This is a very interesting result, because immediately following the hurricane local foresters were concerned that the incredibly devastating effect of the storm on this species might have driven the species locally extinct. Because it is an important species for the local timber industry, such speculation was cause for concern. However, the calculation of $\lambda = 1.156$ showed that not only was the population healthy but one could make a population projection to see that the population will totally dominate the forest by 2014. Obviously such a projection is futile, because our assumption about density independence cannot last forever. Nevertheless, the simple calculation of the dominant eigenvalue put to rest the immediate concerns of local foresters.

From this projection matrix Boucher and Mallona calculated the following survivorship sensitivity matrix:

Stage	Seedling	Small sapling	Large sapling	Small adult	Large adult
Seedling	0.09	0	0	0	0
Small sapling	8.47	0.17	0.04	0.05	0.09
Large sapling	25.11	0.51	0.12	0.16	0.26
Small adult	32.72	0.07	0.15	0.20	0.34
Large adult	40.12	0.82	0.19	0.25	0.42

and the following elasticity matrix:

Stage	Seedling	Small sapling	Large sapling	Small adult	Large adult
Seedling	0.017	0	0	0.017	0.057
Small sapling	0.075	0.098	0.001	0	0
Large sapling	0	0.075	0.041	0	0
Small adult	0	0	0.075	0.128	0
Large adult	0	0	0	0.057	0.358

From the elasticity matrix we see the overwhelming importance of changes in survivorship of the adult trees to changes in λ (the survival of small and large adult trees have the two largest elasticities). This is certainly of concern to foresters as they begin making decisions about harvesting this tree for timber. Note that examining the elasticities does not enable us to answer some potentially interesting questions, because elasticities cannot be calculated for transition probabilities of 0. For example, no adults shrink in size

in the current forest, although saplings may. However, it is perfectly possible that because of the nature of this forest the probability of going from a small adult to a large sapling or from a large adult to a small adult may become important in the future (falling debris becomes a more important factor as the forest matures, meaning that small adults will suffer severe crown damage, effectively turning them into large saplings). The sensitivities can give us some information about the potential impact of this force. The sensitivity for the small adult to the large sapling stage is 0.16 and that for the large adult to the small adult stage 0.34, meaning they have a combined sensitivity (summing the two) of 0.5, which is greater than the survivorship sensitivity of the large adult stage (which, according to the elasticity analysis, was the most important of all).

Population Viability Analysis

One of the most important uses of population projection matrices and related analyses of structured populations is population viability analysis (PVA), whose goal is to estimate probabilities of extinction (usually of rare or endangered taxa), identify the key life history stages or transitions that influence those probabilities, and, often as well, evaluate different management scenarios in terms of reducing the risk of extinction. As you might expect, because we are concerned with probabilities of extinction, PVAs incorporate stochasticity in the demographic functions and so predict population sizes that vary over time. They have become widespread in conservation biology and increasingly sophisticated mathematically (Beissinger and McCullough 2002). Here we provide a simple example of their use for the Florida manatee (*Trichechus manatus latirostris*).

Manatees are listed as endangered and protected by both federal and state law, but nevertheless their major source of mortality is collisions with boats. Marmontel et al. (1997) constructed a projection matrix for Florida manatees using age-specific birth and death rates estimated from 1,212 carcasses of manatees found dead and collected by multiple agencies in the region and then simulated population dynamics under various scenarios. Using 1,000-year-long simulations replicated 100 times for each scenario, they investigated the effects of a number of factors, including environmental variation, periodic catastrophes (e.g., hurricanes, extreme cold fronts, disease outbreaks), and changes in mortality rates to simulate changes in boating accidents. The results are in table 3.3.

The baseline scenario using the carcass-based estimates of demographic variables and no environmental variation or catastrophes (i.e., the deterministic case) had virtually guaranteed the animals' persistence for the 1,000-year period. Adding environmental variation alone increased the risk of extinction only slightly, but adding occasional catastrophes alone decreased the probability of persistence to 61%. Environmental variation and occasional catastrophes together had a synergistic effect, reducing the probability of persistence to 44% and the mean population size an order of magnitude less than

TABLE 3.3. Results of a Population Viability Analysis for Florida Manatees in Different Hypothetical Scenarios Involving Environmental Variation (EV), Periodic Catastrophes, and Mortalities in Different Age Classes (after Marmontel et al. 1997)

Scenario	Mortality	Rate of increase (r)	Probability of persistence	Mean time to first extinction, years	Mean population size
Deterministic	Baseline	0.0004	1.00	None	2,994
EV	Baseline	0.0000	0.94	679	1,250
Catastrophes	Baseline	-0.0030	0.61	843	85
EV + catastrophes	Baseline	-0.0030	0.44	756	161
EV + catastrophes	+ 10% class 0	-0.0070	0.09	606	31
EV + catastrophes	+ 10% class 1	-0.0050	0.15	684	38
EV + catastrophes	+ 10% class 2	-0.0050	0.19	714	52
EV + catastrophes	+ 10% class 3	-0.0050	0.24	675	108
EV + catastrophes	+ 10% class adult	-0.0100	0.00	452	0

Note: Populations were simulated for 1,000 years, starting from an initial size of 2,000 individuals.

in the deterministic case. Note that these fairly drastic changes stem simply from introducing the undoubtedly realistic phenomenon that birth and death rates are not constant in the real world. Keeping these realistic components and adding a relatively small increment of adult mortality of 10% led to a prediction of guaranteed extinction over the 1,000-year period, although this was typically predicted to take more than 400 years. Mortality in younger age classes has somewhat weaker but still considerable effects. Because the great bulk of the mortality is due to boating accidents, these results clearly show that managing the risk of collision is key to maintaining populations of manatees. Marmontel et al. also varied the carrying capacity of the population they studied, simulating changes in food availability; this had relatively little effect on the population dynamics.

Demography of Invasive and Native Plant Populations

The tools of elasticity and sensitivity analysis are useful not only for investigating specific systems as in the examples above, but also, and even more so, for comparative analyses that enable syntheses of patterns of life history and demography across many species and environments. Because elasticity, in particular, is scaled to represent the proportional contribution of different stages to population growth and is therefore additive across stages (review chapter 2), it can be very straightforward to compare the life histories of quite different species living in different environments. In the past decade, metaanalyses synthesizing elasticities or related metrics across multiple studies to answer basic questions in life history have become increasingly common, and some interesting patterns have emerged. Many of these analyses have followed Franco and Silvertown (1996) in classifying transitions into three general categories: reproduction (the first row of a projection matrix), survival and regression (the diagonal and supradiagonal of a projection matrix, i.e., the probability of staying at the same stage or getting smaller), and growth (the subdiagonal of a projection matrix, i.e., the probability of advancing to a later or larger stage) (see figure 2.3). Using this framework, syntheses in both plants (Franco and Silvertown 2004) and mammals (Oli and Dobson 2003) have shown that the relative importance of fecundity versus survival in determining per capita population growth depends on life span. Shorter-lived species tend to respond more to changes in fecundity, and longer-lived species respond almost entirely to changes in survival, with very little impact from changes in fecundity or growth.

A recent example elaborates these patterns, comparing the demography and population dynamics of 21 plant species where they were invasive and 179 plant species in their native habitat (Ramula et al. 2008). Using the standard approaches to determining the asymptotic λ , Ramula and associates found that λ was significantly higher for the invasives ($\lambda = 1.47$) than for the native taxa ($\lambda = 1.05$). These values indicate that native species are nearly stable but that invasives are rapidly increasing—not surprising, perhaps, but suggesting that the invaders studied are all still in the exponential

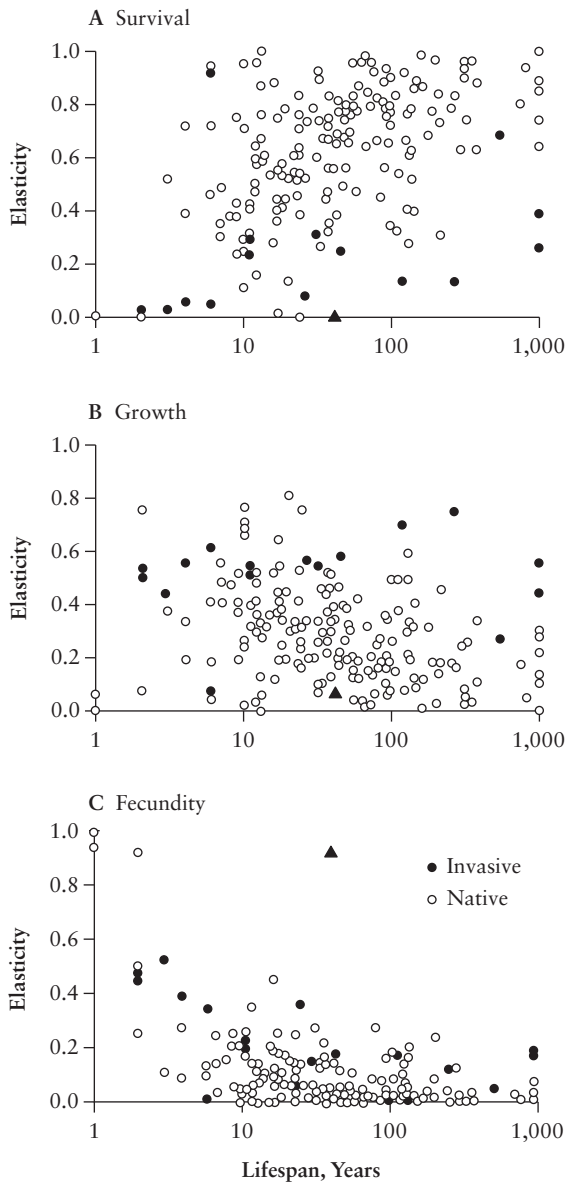


FIGURE 3.3. Elasticity of survival, growth, and fecundity for 29 invasive and 179 native species as a function of life span (from Ramula et al. 2008).

phase of their expansion. As has been found in other studies, elasticity of survival increases with life span, while elasticity of fecundity declines with life span (figure 3.3). Figure 3.3 also illustrates that although this is true for both invaders and natives, invaders consistently have greater elasticity of fecundity and growth but lower elasticity of survival than do natives. That is, not only are the population dynamics different but the important life his-

tory drivers also differ between invaders and natives. These results could be quite important to management. For example, using simulations to confirm the results of the elasticity analysis, Ramula et al. further suggested that targeting control methods to reduce growth and fecundity would be most effective in managing a short-lived invader, but methods to target survival and growth would be more effective for a long-lived invader.

A Closer Look at the “Dynamics” in Population Dynamics

4

Central to the analysis of population dynamics are concepts that come naturally to anyone trained in the physical sciences. In elementary physics classes, for example, a physical system is most frequently looked at from the point of view of stability and equilibrium. When engineers design systems, in fields from aerospace to industrial management, one of the first questions asked is, Under what conditions will the system be at equilibrium, and will it be stable or not? Ecologists also began by asking such questions of ecosystems. As a consequence, concepts such as balance (equilibrium) and stability have become central to both mathematical and conceptual analyses of all ecosystems, especially in population and community ecology. Often such concepts have even become normative, central goals toward which we must strive in the design of sustainable resource management systems. Some ecosystems are reportedly unstable and have lost the inherent equilibria of natural ecosystems (Altieri 1987; Soule et al. 1990), and the job of good husbandry should be to promote the use of management tools that will restore balance and stability, and therefore sustainability, to the system (Altieri 1987; Levins and Vandermeer 1989). Such is a widely held position.

In population and community ecology, such ideas have been debated and clarified over the past 20 years and their meanings operationalized to a considerable extent. Furthermore, this operationalization forced a rethinking of some central concepts. To take a concrete example, one of the ideas that has seen a great deal of rethinking is that the complexity of a system is related to its stability. Metaphorically, as a system becomes larger, it develops more interconnections and, much as in the case of a spider web, those interconnections make it resilient to outside perturbations and thus stable—the more interconnections (read, the larger and more complex the system), the more stable it should be. However, the notion that a large, highly connected system would be more stable than a small system with low connectivity was challenged by

May (1974), who proved that, all else being equal, the larger and more complex the system, the more likely it is to be unstable, precisely the opposite of what most ecologists intuitively felt. Rather than being like a spider web that receives its stability from all of the interconnections within it, food webs appeared to be like houses of cards, deriving their structure from the myriad connections among parts but becoming more fragile the larger they are.

The original intuition of ecologists was that large ecosystems, with their great biodiversity and complex interconnections, are more stable and more in “balance” (at equilibrium) than are simplified systems that have purposefully been designed to eliminate much of that complexity. It seems an obvious idea, so how can it be that careful analytical thought suggests otherwise? One of the problems, perhaps the principal problem, is that these early conceptualizations were based on classical notions of dynamics (stability, equilibrium, and balance) from the physical sciences. With those notions coupled with the semiromantic notion of the balance of nature, a mainstay of nature lovers and environmental activists alike, a perceived concatenation of the popular with the scientific created what seemed to be an unassailable principle. But a false analogy had been built up between the naturalists’ notion of balance and stability and the classical engineer’s notion of those same things. With more modern interpretations of the underlying dynamic structure of these and other concepts, largely derived from the new science of nonlinear dynamic systems, a new classification of dynamic behaviors may better correspond to the old naturalists’ or traditional farmers’ original intuition (Vandermeer et al. 2010). Appreciation of this idea will be enhanced by the material offered in this chapter.

EXERCISES

4.1 The integrated form of the logistic equation is

$$N_t = \frac{KN_0}{(K - N_0)e^{-rt} + N_0}.$$

Use this equation to generate a time series of N versus t . Use parameters $r = 1.5$, $K = 100$ and construct two time series, one with $N_0 = 10$ and one with $N_0 = 190$. (Use a time frame from 1 to 5 with intervals of 0.1.)

4.2 Repeat exercise 4.1 with $r = -0.1$. (Use a time frame from 0 to 100 with intervals of at least 1.0.)

4.3 Recall the population model of the Ricker equation from chapter 1,

$$N_{t+1} = rN_t e^{(1-bN_t)}.$$

Let $r = 2.5$ and $b = 2.5$ and project the population 45 time units, beginning with a starting population of 0.01. What is the pattern of the time series? Let $r = 4$ and $b = 0.7$ and project the time series 45 times, beginning with a starting population of 3.5. Now what is the pattern of the time series?

- 4.4 A simple model of predator (P) and prey (V , for victim) interacting in discrete time is, for the prey,

$$V_{t+1} = bV_t \left(\frac{K - V_t}{K} \right) e^{-aP_t}$$

and for the predator,

$$P_{t+1} = cV_t(1 - e^{-aP_t}).$$

This model will be developed more fully in chapter 6. For now just examine the time series it generates. Use the parameters $a = 0.1$, $b = 1.5$, $K = 40$, and $c = 1.5$. Make a graph of the two species over time (from 0 to 200), and plot the two species on a graph of predator versus prey, showing the direction of change with arrows.

- 4.5 An even simpler model (again, to be developed in chapter 6) is

$$\begin{aligned} V_{t+1} &= bV_t - mV_tP_t, \\ P_{t+1} &= b'V_tP_t - m'P_t. \end{aligned}$$

Repeat exercise 4.4 with this model and the parameters $b = 1.1$, $m = 0.8$, $b' = 0.5$, and $m' = 0.002$.

Intuitive Ideas of Equilibrium and Stability

The intuitive notions of balance and stability have their parallels in classical analytical thought, balance as equilibrium and stability as one of two forms, either unstable or stable. Consider the graph in figure 4.1. The variable x could be any interesting variable, but for our purposes it is best to think of it as population density. Plotting density over time, beginning at various starting points, we see that no matter where the trajectory begins, it always ends up at the value x^* . Furthermore, once it attains the value of x^* it never deviates. The value x^* is thus an equilibrium point (the system is in “balance” once it reaches that point). However, that the system is in balance is only one feature of x^* that is important. The behavior of the variable x when it is not exactly at that equilibrium point is also of great importance. Although it is true that when the system is at equilibrium (the variable x exactly equals x^*) it will remain there in perpetuity, it is also true that the slightest deviation from that value (say, to the point labeled “deviation from equilibrium” in figure 4.1) will result in a return to that same equilibrium. Because any such deviation will result in a return to the same equilibrium point, the point is a “stable” equilibrium point.

In contrast, consider the situation presented in figure 4.2. Again there is an equilibrium point (x^*), and again, if the system is initiated at exactly that point, it will remain there in perpetuity. But the slightest deviation from that point means the system will deviate forever. Thus the point is in balance and it is an equilibrium point because it will stay where it is forever if undisturbed. However, in this case the slightest deviation results in continued deviation. This is referred to as an unstable equilibrium point.

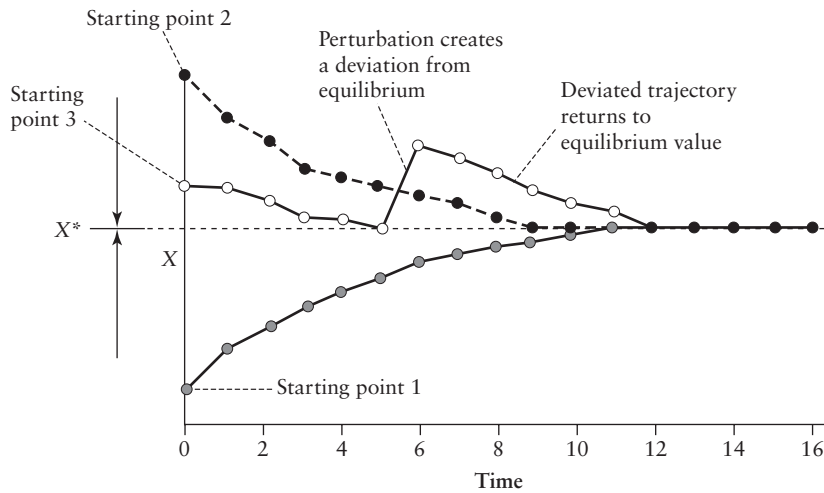


FIGURE 4.1. Illustration of the dynamical behavior associated with a point attractor.

In the past, equilibrium points have been called fixed points, singularities, and probably other things as well. The adjective *stable* or *unstable* is then attached to indicate the dynamic behavior of points near that singularity (or fixed point or equilibrium point). In more recent literature, the notion of the equilibrium itself and the behavior of points near to it has been termed either an attractor (for a stable equilibrium point) or a repeller (for an unstable equilibrium point). The terms *attractor* and *repeller* are more suitable for discussion given the recent advances in our understanding of models that have

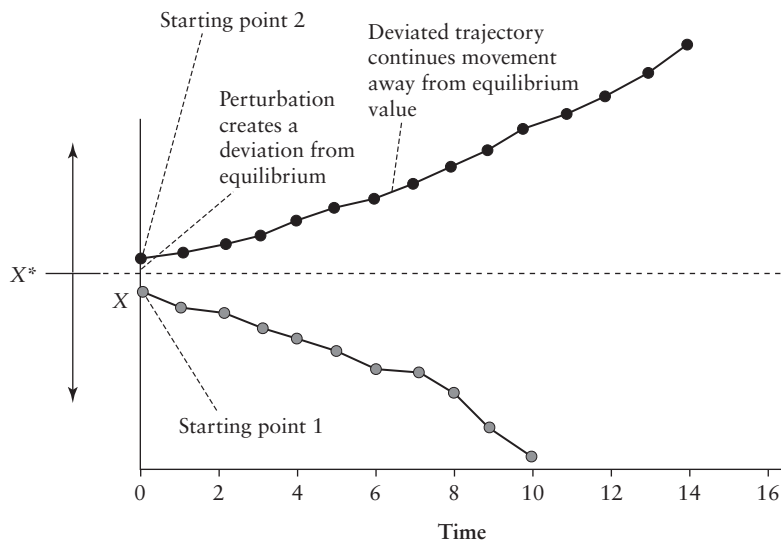


FIGURE 4.2. Illustration of the dynamical behavior of a point repeller.

the sort of complexity demanded by ecological systems. In the rest of this text, when the subject arises the terms *attractor* and *repeller* will be used rather than *stable equilibrium* and *unstable equilibrium*.

To rapidly picture the dynamics of a system it is often convenient to simply represent the attractor or the repeller as a point on the line that represents the possible range of the variable in question (the ordinate in figures 4.1 and 4.2), with small arrows indicating the direction in which the trajectories near to the equilibrium point will go. The line is called the state space (i.e., the space, in the mathematical sense, that represents all possible values or “states” of these variables, called state variables), and the collection of arrows is called the vector field. When the arrows point toward the equilibrium point (as in figure 4.1) it is an attractor, and when the arrows point away from the equilibrium point (as in figure 4.2) it is a repeller. Thus an examination of the vector field reveals whether a point is an attractor or a repeller or something else.

It is also popular to indicate the dynamics of a system by means of small physical models, as in figure 4.3. The marble on top of the hill (figure 4.3A) illustrates a repeller (the line below it with the point and the arrows is equivalent to the ordinate of figure 4.2 turned on its side), and the marble at the bottom of the valley (figure 4.3B) illustrates an attractor. Because the attractor and the repeller are single points, they are called a point attractor and a point repeller.

Another major category of behavior is not representable in such simple diagrams but requires a two-dimensional space (three-dimensional including time). Suppose we have a beaker of water whose bottom has the positive end of a magnet affixed to its center. We then drop a smaller magnet into the beaker with its negative pole facing downward and watch what it does as it falls through the water. In figure 4.4A the mobile magnet falls toward the magnet on the bottom. If it is placed in the water at exactly the center of the beaker, it will remain in this position (actually somewhere above this position) as it falls

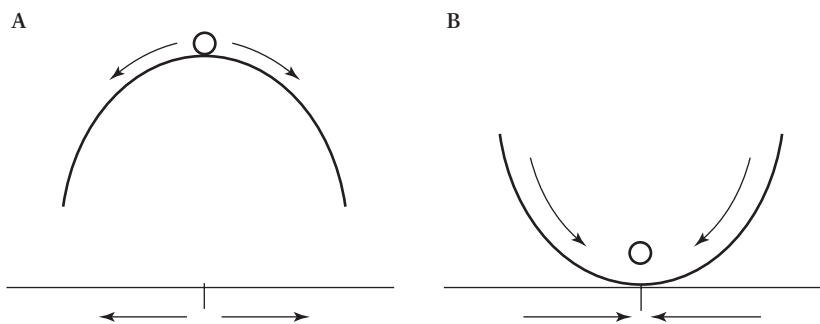


FIGURE 4.3. Physical models of a classical attractor and repeller. (A) The marble is balanced on top of the hill at equilibrium, but the slightest deviation from that point results in continued deviation, corresponding to the situation in figure 4.2. (B) The marble is located in the valley, and all deviations from that point result in a return to it, corresponding to the situation in figure 4.1.

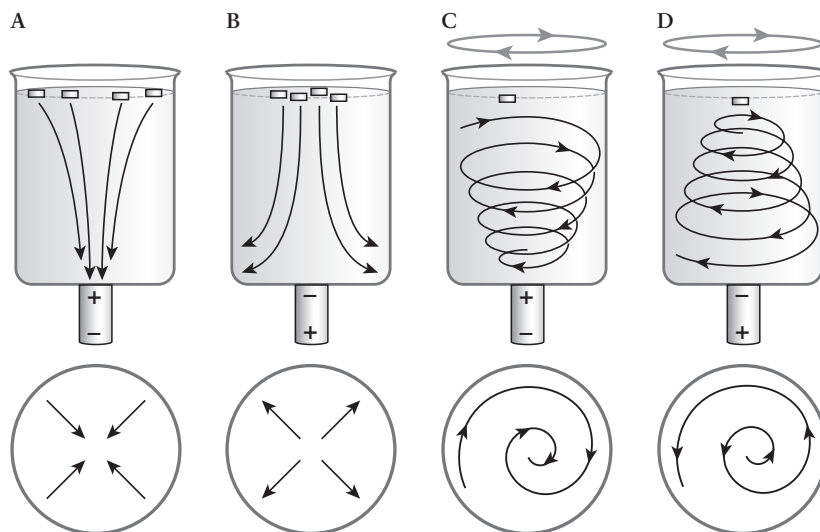


FIGURE 4.4. Beaker and magnet model of dynamics. (A) The small magnets in the water, with their negative poles pointed downward, are attracted to the positive pole of the magnet on the bottom of the beaker. (B) The small magnets are repelled from the negative pole on the bottom of the beaker. (C) When the beaker is constantly rotated, the magnet undergoes a spiraling motion as it descends through the water toward the positive pole of the magnet on the bottom. (D) When the beaker is constantly rotated, the magnet undergoes a spiraling motion as it descends through the water away from the negative pole of the magnet on the bottom. The circle at the bottom of each diagram illustrates the general behavior of the small magnet as viewed from the top (or bottom) of the beaker.

through the water. If it is placed somewhere deviant from this position, it will fall toward the bottom magnet. Obviously this physical model is identical to the example in figure 4.1 except that it is in three dimensions, the horizontal two dimensions of the bottom or top of the beaker and the vertical dimension that represents time (the position from the top of the beaker to the bottom is proportional to the time since the small magnet was dropped into the beaker). Just as we could represent the behavior of the system on a line (the ordinates in figures 4.1 and 4.2 and the lines below the diagrams in figure 4.3), we can do so by looking at just the bottom of the beaker, as shown in the circle below each diagram of the beaker in figure 4.4. Figure 4.4A represents an attractor, and figure 4.4B represents a repeller.

With the beaker model we can see another class of behavior that is extremely important in ecological models. Suppose the beaker is placed on a mixing table that creates a vortex in the water. The expectation is that whatever is dropped into the beaker will spiral around as it drops through the water, as indicated in figure 4.4C. However, as it spirals around it is also attracted by the magnet on the bottom of the beaker. For obvious reasons this attractor is referred to as an oscillatory point attractor. The parallel behavior of the repelling magnet in swirling water is also oscillatory, but it is a

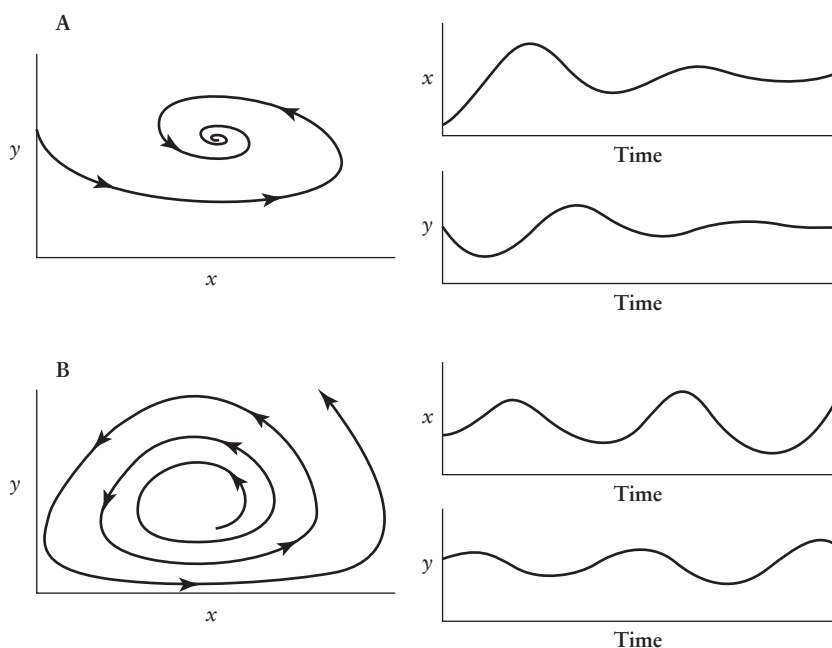


FIGURE 4.5. Traditional representations of an oscillatory attractor (A) and an oscillatory repeller (B). x represents prey, and y represents predator. The graph of y versus x is the traditional “phase plane” diagram. The same data are plotted to the right as a time series in both variables.

repeller—an oscillatory point repeller (figure 4.4D). For each of the oscillatory points (attractor and repeller) the picture on the bottom of the beaker is a spiral (figure 4.4C and D).

Figure 4.5 illustrates how this spiraling behavior looks in a more traditional diagram of the variables over time. The two dimensions of the beaker’s bottom are plotted over time to illustrate that an oscillatory attractor is the same as “damped” oscillations, whereas an oscillatory repeller is the same as expanding oscillations.

Yet another class of behavior is extremely important in physical as well as biological systems. This class requires a different physical model, as illustrated in figure 4.6.A. A small hill in the middle of a valley causes a marble to roll down the hill but to become entrapped in the valley, rolling continuously around the bottom of the valley. The sides of the valley cause any marble beginning on that surface to wind down the valley floor, again rolling around the bottom of the valley. The ultimate fate of any trajectory is either to move to the outer limits of the hill or to wind up cycling forever in the bottom of the circular valley (presuming that there is some sort of energy that keeps the system in motion). This kind of behavior is known as a periodic attractor, so named because at some time in the future the system always returns to the same position, which is to say that it periodically returns to any given state. Note that the example in figure 4.6A actually includes two periodic cycles,

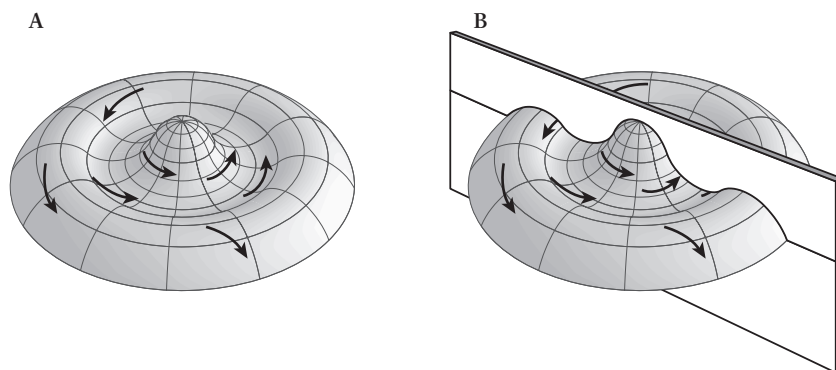


FIGURE 4.6. Physical model illustrating a periodic attractor (limit cycle) (see text).

an obvious one at the bottom of the valley and a not-so-obvious one exactly on the outer edge of the valley. That is, it is theoretically possible to have a marble cycling around the top of the outer boundaries of the valley, always exactly balanced between the force attracting it down to the valley bottom and the force attracting it down off of the side of the hill. Obviously this limit cycle cannot really be observed because the marble cannot be expected to maintain exactly this balance. It is thus an unstable cycle or a periodic repeller.

Taking a cross section of this model (figure 4.6B), we arrive at the more easily interpretable section pictured in figure 4.7 (such a section is formally a Poincaré section). As before, we can summarize the overall behavior of the system with little arrows on the line. Here we see three repellers and two attractors, although the two attractors are simply two points on the periodic

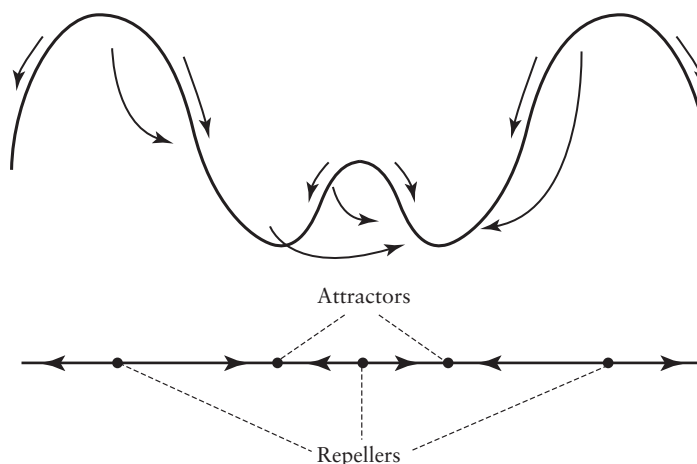


FIGURE 4.7. Cross section (Poincaré section) through the surface of figure 4.6, showing how the dynamics of the system can be illustrated.

attractor and the two outer repellers are simply two points on the periodic repeller. Thus they are not really equilibrium points, as in previous examples, but rather points on an attractor (or points on a repeller).

The final type of qualitatively distinct behavior that is commonly observed in ecological models results if we assume that the bottom of the valley is perfectly flat. If the beaker model of figure 4.4 has the magnets removed from the base or if the hill model of figure 4.6 has its bottom constructed to be absolutely flat, the physical attraction (the magnet in the beaker model, the force of gravity in the hill model) has been removed and we theoretically expect the system to move around in this space, constrained to be sure but without a tendency to move to the center within that space, as suggested by the model in figure 4.8. As before, there are three repellers but no attractors, at least not of the sort in previous examples. Yet the entire bottom of the valley will certainly attract the marble, and in this intuitive sense it is also an attractor. But here we have an attractor that is neither a point nor a cycle but rather an area or region. Being a region that attracts all trajectories yet has no tendency within it to move to the center (no point attractor), it is thought to be rather strange. This is why it is referred to as a strange attractor, and the behavior of a system within it is referred to as chaotic.

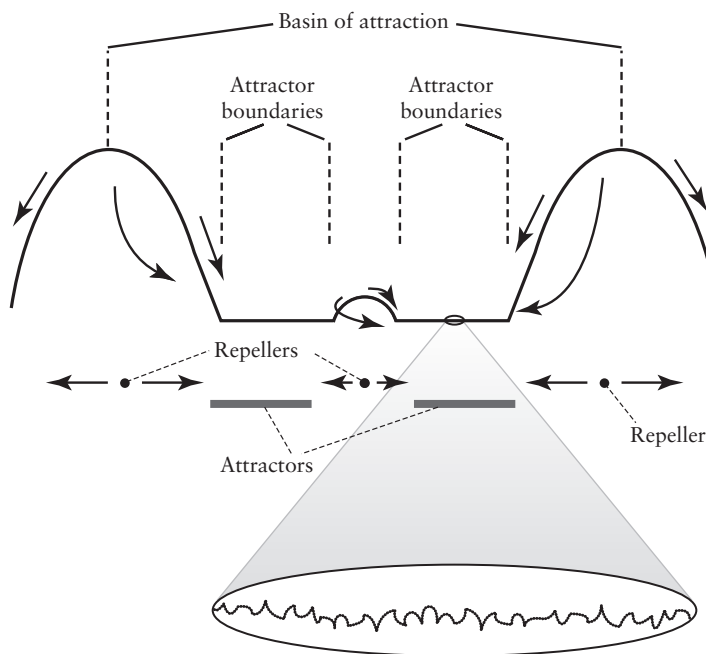


FIGURE 4.8. Poincaré section similar to that in figure 4.7 but with a strange attractor rather than a periodic attractor. The bottom of the valley is, theoretically, perfectly flat, so there is no natural place to which the marble will be attracted. So the entire flat region will attract the marble because the walls of the valley still slope downward. But once it reaches the floor of the valley, its motion will become unpredictable. This is a strange attractor.

Obviously there is a qualitative difference between this type of attractor and those discussed previously. There is no particular point to which the system ultimately tends but rather an area to which it tends. Furthermore, from a practical standpoint we cannot be interested in the final state of the system because it has multiple final states (in the sense of a single point). The concern really ought to be with the range within which the system will ultimately be found, as discussed below.

The class of behaviors illustrated by the simple hills (figure 4.3 and 4.6) or the beakers (figure 4.4) are the classical behaviors usually analyzed by engineers. Certainly they are also important points of departure for analyzing ecological systems. However, there are other kinds of behaviors, most importantly those illustrated in figure 4.8, in which the focus is on the range of expected values and the persistent changes through time within that range. In summary, attractors (or repellers) can be thought of as falling on a gradient going from simple point attractors (or stable equilibria, stable nodes, stable fixed points—all synonyms—as in figure 4.3) to oscillatory attractors (or stable foci—still point attractors—as in figure 4.4C) to periodic attractors (also called limit cycles) to strange attractors (or chaotic attractors, as in figure 4.8; the subject of chaos will be discussed later in this chapter).

One's interest in analyzing a system depends on the nature of the equilibrium state. If a point equilibrium exists, for example, a central question is how to locate the exact position of the point and determine whether it is an attractor or a repeller. This is the focus of the classical engineering sciences. But if a strange attractor exists, the interest is more in locating the position of its boundaries and discovering other details about its "morphology," as discussed later.

One further concept is especially important when dealing with strange attractors. The "basin of attraction" is the collection of the values of the state variables from which all trajectories eventually wind up exactly on the attractor. In figure 4.7, for example, the tops of the two largest hills represent the outer edges of the basin of attraction for the limit cycle attractor at the bottom of the valley, and the small hill in the middle represents the inner edge of that basin. The edges of a basin of attraction are always repellers, as is evident in figure 4.7. The edges of the basin of attraction are not the same as the boundaries of a strange attractor. The latter refer to the outer limits that the attractor itself can realize, the former to all possible states that eventually reach the attractor. Formally speaking, the attractor (and its boundaries) is a subset of the basin of attraction but not the reverse (see, for example, figure 4.8).

Classical ecological theory has dealt mainly with point attractors and to some extent with periodic attractors. Only with the advent of nonlinear dynamics as a theoretical science has there been a realization that the alternative type of equilibrium and stability actually exists, that is, the strange attractor. Such attractors become more common in the literature as old models are analyzed more completely and especially as new model situations are explored. These attractors have also received considerable attention simply because they are sometimes called chaos or chaotic attractors. This unfortunate choice of terminology will be further discussed later in this chapter.

For now suffice it to say that a strange attractor, because of its so-called chaotic motion, is unpredictable in a very special technical sense. This fact has caused considerable unnecessary consternation among those who seek to predict natural phenomena and has led to a small cottage industry of researchers attempting to show that particular data sets do or do not represent true chaotic behavior. However, the importance of the issue lies not with the distinction between chaos and nonchaos, despite what popular articles contend, but rather with the distinction among point, periodic, and strange attractors, three positions on a continuum from point to strange, as discussed above. In the case of point attractors we are concerned with the location of the equilibrium and its stability properties. In the case of strange attractors we are concerned with their boundaries and qualitative behaviors, their morphology.

EXERCISES

- 4.6** The basic exponential equation describes the dynamics of a single population, which means that it is a dynamic system in one dimension. Its only equilibrium value is $N^* = 0$. Draw the state space (a line), showing with small arrows the dynamical nature of the system near the equilibrium point. Make a graph of the derivative (dN/dt) as a function of N for $r = 1.0$ and 1.5 .
- 4.7** The logistic equation also describes the dynamics of a single population but with two equilibrium points, K and 0 . Draw the state space showing the dynamical nature of the system near the equilibrium points. Make a graph of the derivative (dN/dt) as a function of N for $r = 1.0$, $K = 1.2$; for $r = 1.5$, $K = 1.2$; and for $r = 1.5$, $K = 1.8$.
- 4.8** If you have a single population model based on a single well-behaved ordinary differential equation and it has five equilibrium points and diverges to infinity at very large values, what must the collection of vectors (formally called the vector field) look like (again, on a single line, the relevant state space)? Sketch what you think a graph of dN/dt versus N would look like.
- 4.9** Assume that a population is growing according to the logistic equation. To make things simple, presume that the value of both r and K is 1.0 (i.e., we represent the population as varying between 0 and 1). The equilibrium value of that population will be

$$0 = N - N^2,$$

which is a quadratic equation and has two roots, which are, by inspection, 0 and 1 . Now suppose that a manager decides to impose a fixed harvesting rate on the population such that a constant number of individuals will be removed each year. A sensible model for this situation would be

$$\frac{dN}{dt} = N(1 - N) - N_F,$$

where N_F is the fixed number (actually the proportion) of individuals removed. Plot the derivative versus N for the following values of N_F : 0 , 0.25 , and 0.5 . Also, directly solve for the roots (using the quadratic formula; remember, N_F is a constant). What do you conclude?

Eigenvalues: A Key Concept in Dynamic Analysis

Consider a simple point attractor in one dimension. As above, we can represent its qualitative dynamics by drawing the state space (a line) and indicating where the point is in that space (on that line) and then adding the vector field (small arrows indicating the direction and rate of change), as presented in figure 4.9.

If we now rotate the vectors 90 degrees, either upward to illustrate an increasing vector or downward to illustrate a decreasing vector, we get the picture shown in figure 4.10. If we connect the tips of the rotated arrows with a line, the slope of that line is called the eigenvalue. With this formulation (figure 4.10) it is evident what the eigenvalue means in this case; it is the rate at which the system approaches a point attractor (or leaves a point repeller, if the rotated arrows go in the opposite direction).

To relate the general concept of eigenvalues to population-dynamic models, recall the exponential equation from chapter 1,

$$\frac{dN}{dt} = rN.$$

This is a system of one dimension (a single variable, N), and thus its state space and dynamics are as in figure 4.9. The equilibrium point is at $N = 0$, so the left side of the state space does not exist for this model. Suppose that $r < 0$,

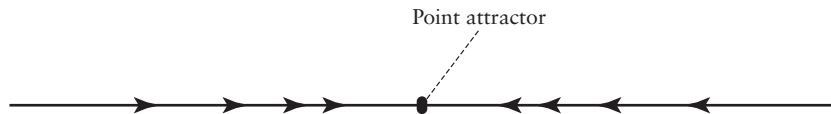


FIGURE 4.9. State space for a one-dimensional (one-variable) model, illustrating a single point attractor and its vector field (the collection of arrows indicating the dynamics of the system).

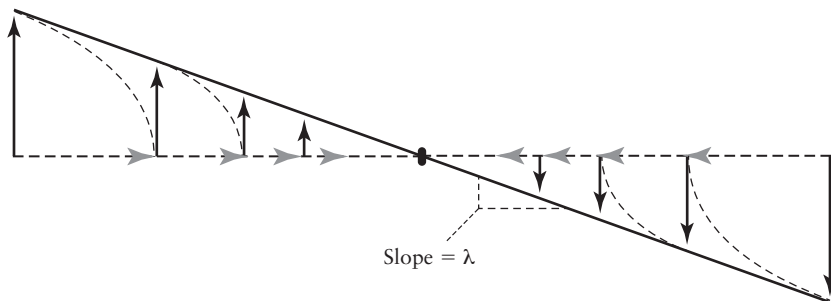


FIGURE 4.10. The vectors of the example from figure 4.9 rotated. The vectors to the right of the point attractor are decreasing, so we rotate them downward (decreasing). The vectors to the left of the point attractor are increasing, so we rotate them upward (increasing). Then we connect the arrowheads with a line. The slope of the line is the eigenvalue of the point attractor.

which is to say, a declining population, one that will eventually go locally extinct. A plot of dN/dt versus N will look something like the right side of figure 4.10 with the slope $= r$. That is, because the vectors represent dN/dt at particular values of N , rotating them 90 degrees is the same as plotting them on the y axis. Thus we see that, in this example, the eigenvalue of the attractor is r (because a plot of dN/dt against N is linear in the case of the exponential equation).

If the population is growing ($r > 0$), the result is qualitatively different in that the arrows will all be pointing away from the equilibrium point, that is, the point is a repeller (the equilibrium point is still zero). Thus the arrows to the right of the point will be rotated clockwise (the opposite of what we see in figure 4.10), and again the part of the graph to the left of the point does not exist for this model. The line connecting the arrowheads will thus have a positive slope, which means a positive eigenvalue.

Now, suppose that we have a population growing according to the logistic equation. Its dynamics (again in one dimension) will look something like what is pictured in figure 4.11.

Here we have two equilibrium points, one an attractor at the carrying capacity and one a repeller at the value of zero. If we now rotate the arrows, as before, we obtain the graph shown in figure 4.12.

Here there is no simple slope to the line, but in the neighborhood of each of the equilibrium points we can approximate the curve with a straight line, and the slope of that straight line is the eigenvalue associated with the equi-



FIGURE 4.11. State space for a one-dimensional model based on the logistic equation. There are two equilibrium points, one an attractor (K), the other a repeller (0).

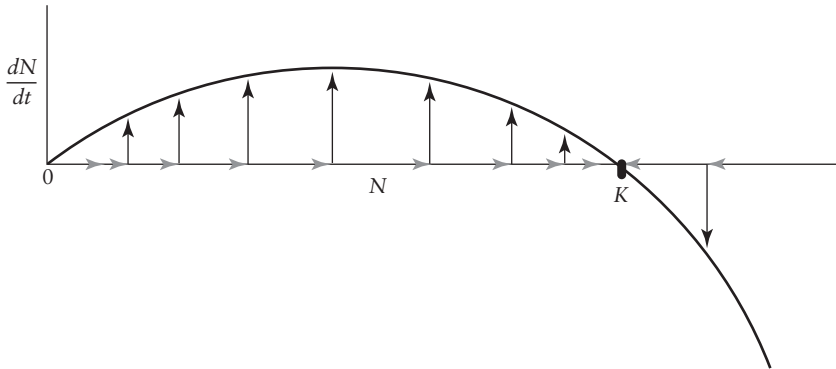


FIGURE 4.12. Dynamics of the logistic equation in one dimension, with the changes in the derivative graphed as the ordinate.

librium point. For example, consider the equilibrium point at the carrying capacity (K). The slope of the curve at that point is simply the derivative with respect to N of the derivative with respect to time, evaluated at K , or

$$\frac{d\left(\frac{dN}{dt}\right)}{dN} = r - \frac{2rN}{K}.$$

Substituting K for N (because the slope is the derivative evaluated at K), we obtain

$$\left.\frac{d\left(\frac{dN}{dt}\right)}{dN}\right|_{N=K} = r - \frac{2rK}{K} = r - 2r = -r.$$

which tells us, first, that the point K is an attractor (because the eigenvalue, $-r$, is negative) and second, that the rate of approach to that equilibrium will be $-r$. A similar procedure applied at the other equilibrium point (0) gives an eigenvalue of r , showing that it is a repeller (because r is positive) and that the rate of deviation from it is r .

Note that the eigenvalues computed for the projection matrices of the previous chapter have precisely the same qualitative meaning as in the present chapter. However, earlier we discussed only the computation of eigenvalues for a matrix with constant values, in which case the population was always an exponential population with a single equilibrium point at zero. If the population was growing, its largest eigenvalue was positive and it was growing at a rate equal to the value of that eigenvalue. A negative dominant (largest) eigenvalue indicated, as it does here, that the equilibrium point is an attractor, which means that the population is declining, and the rate of that decline is the value of that eigenvalue. So we see that the dominant eigenvalue of a projection matrix (without density dependence) is precisely the same as the eigenvalue of the exponential equation, r .

In two or more dimensions (i.e., when we have two or more species interacting, so two or more state variables), the situation is a bit more complicated. In two dimensions the state space is the plane, and we must examine the dynamics of the system in that plane in a third dimension. We have already looked at this issue in a very qualitative way in figure 4.4 (the first two panels, representing the stationary beaker), where the fixed magnet either attracted or repelled the falling magnet. We now examine the two-dimensional case in more detail. Consider the physical model in figure 4.13. A marble rolling on this surface will eventually wind up at the point where the two folds intersect, but there will be a bias in that most of the time it will roll down along the fold labeled A. It would be possible for it to roll directly down fold B and arrive at the equilibrium point, but this would be very unlikely because that fold is a knife edge on which the marble would have to balance as it rolled down.

For heuristic purposes it makes sense to ask what would happen if the marble began exactly on the fold A. Now we can represent the system in a single dimension, a dimension along fold A, and look at the dynamics along

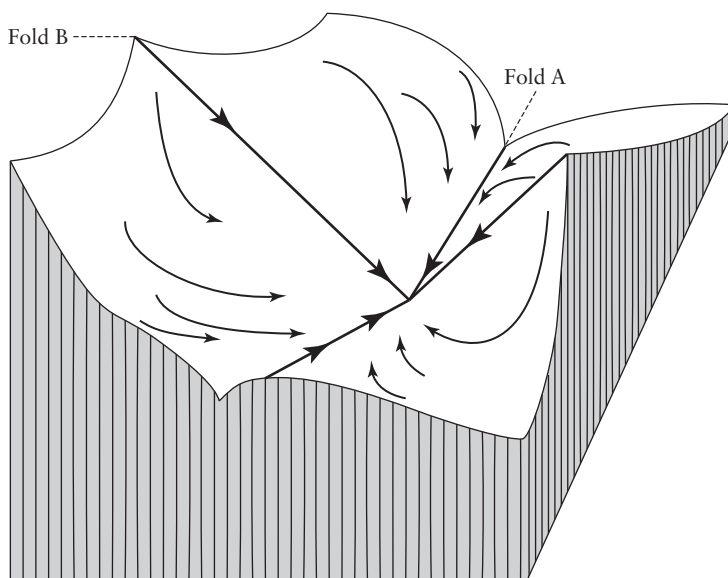


FIGURE 4.13. Physical model of the dynamics of a point attractor in two dimensions.

this fold as if it were a one-dimensional system. Thus the analysis reverts exactly to the one-dimensional analysis we did in figures 4.9–4.12 above. And indeed the rate of change of the rate of change along this fold is an eigenvalue.

But there is still the theoretical possibility that the marble will balance on fold B and, like a tightrope walker, roll down, precariously balanced, until it reaches the equilibrium point. As unlikely as that may seem, we can still analyze it mathematically using the graphical method we used in figures 4.9–4.12. Again we come up with a measure of the rate of change of the rate of change as we approach the equilibrium point, and that rate is an eigenvalue. Thus we see how, when we have two dimensions, we have two eigenvalues. Indeed it is the case that there will always be as many eigenvalues as there are dimensions in the system.

Here we see the significance of the “dominant” eigenvalue. It is the value of the rate of change of the rate of change along the dominant fold (fold A in figure 4.13), that is, the rate at which the system will approach the equilibrium point as it gets close to it. There will always be one collection of points (a “fold”) along which the marble will eventually tend, and that collection of points defines the one-dimensional system that is used to calculate the dominant eigenvalue (see figures 4.9 and 4.10).

In figure 4.14 the three possible configurations in two dimensions are illustrated, along with the eigenvalue states, a point attractor when both eigenvalues are negative, a point repeller when all the eigenvalues are positive, and a “saddle” point when one eigenvalue is negative and the other positive. Clearly, an examination of the signs of the eigenvalues provides a definitive statement as to which of the situations exists. Two positive eigenvalues indicate a simple

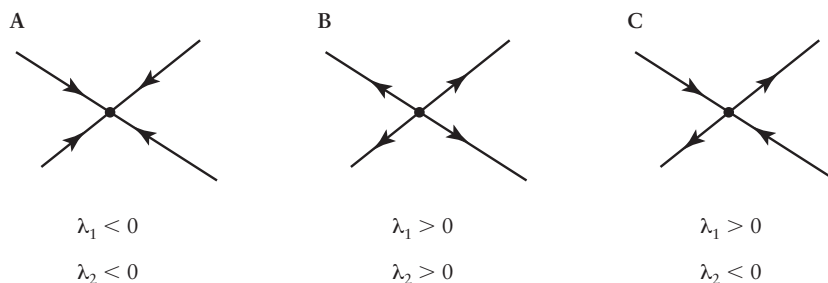


FIGURE 4.14. Conditions of eigenvalues for the three most common qualitatively distinct arrangements in two dimensions. (A) Point attractor. (B) Point repeller. (C) Saddle point repeller.

point repeller, two negative eigenvalues indicate a simple point attractor, and a positive and a negative eigenvalue indicate a different kind of repeller. Note that in the latter case the point is approached from some lines and repelled along other lines, much as a marble would be when rolling along the surface of a saddle. For this reason, this sort of equilibrium is referred to as a saddle point repeller.

So far our presentation has been largely graphical and heuristic. In reality, for a given model simple recipes exist for finding the eigenvalues of a system at a point (indeed, in the contemporary world a few keystrokes or pointing and clicking on the “find eigenvalue” button is usually the way to do it). Frequently the eigenvalues turn out to be simple real numbers and one merely has to compare them to zero to determine the qualitative nature of the point. But sometimes they turn out to be complex numbers, that is,

$$\lambda = r + ci,$$

where i is the square root of -1 . Thus there is a real part (r) and an imaginary part (c). There is no convenient way of explaining exactly why, but the fact is that oscillatory systems (e.g., the swirling beaker model of figure 4.4C, D) have eigenvalues with nonzero imaginary parts. The parallel graphs of the ones already made in figure 4.14 are shown in figure 4.15 for oscilla-

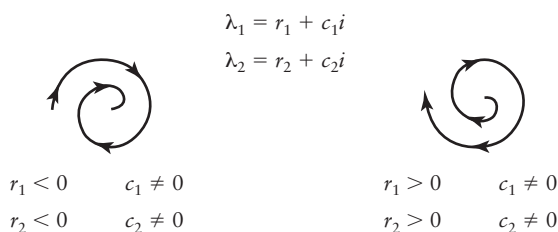


FIGURE 4.15. Conditions of eigenvalues for the two most common qualitatively distinct arrangements in two dimensions when the eigenvalues have nonzero imaginary parts.

tory point attractors and repellers. It is a simple rule that nonzero imaginary parts of the eigenvalues mean that the system is oscillatory, and the oscillations wind down to the equilibrium point if the real values are negative and wind away from the point if the real values are positive.

EXERCISES

- 4.10** In chapter 1 (exercise 1.17) you used the logistic map to project the population 50 times with values of $\lambda = 1.5, 2.0, 3.0$, and 3.5 . Do the same for $\lambda = 3.4, 3.5, 3.56, 3.565$, and 3.567 , starting with 0.8 individuals and projecting 100 time steps. Note that at the end of the run the numbers tend to repeat themselves in a regular sequence. For example, for $\lambda = 3.4$ the numbers go from 0.452 to 0.842 and then back again to 0.452, which is to say that there is a two-point cycle. How many points are in the cycles at the ends of the runs of the other values of λ ? Plot the numbers in a particular cycle versus the value of λ that gave those numbers.
- 4.11** Add to the graph of 4.10 the cycles for $\lambda = 3.43, 3.455$, and 3.53 .
- 4.12** Add to the graph of 4.11 the cycle (approximate cycle) for $\lambda = 3.6$.
-

Basic Concepts of Equilibrium and Stability in One-Dimensional Maps

So far we have focused on models in continuous time and, although we didn't show them explicitly, these are based on differential equations. Although the same basic dynamical concepts apply to models in discrete time using difference equations, the development and rules are actually quite different. Let us suppose that a reasonable model of population dynamics is a mapping from one time period to the next, which is to say that the population density at this point in time is some multiple of what it was in the previous time period. That is, if N_t is the population density at time t , we have, for example,

$$N_{t+1} = \lambda N_t, \quad (1)$$

where λ is, at this point, an arbitrary constant. This is an alternative way of expressing the exponential growth of a population. The relationship between this form and the differential equation form of chapter 1 is as follows. From the equation

$$\frac{dN}{dt} = rN,$$

integrate to obtain

$$N_t = N_0 e^{rt}, \quad (2)$$

which means we can also write

$$N_{t+1} = N_0 e^{r(t+1)} = N_0 e^{rt} e^r.$$

Substituting from equation 2, we have

$$N_{t+1} = N_t e^r,$$

and letting $\lambda = e^r$ we substitute to obtain equation 1, making it obvious that the one-dimensional map is exactly equivalent to the more traditional differential equation. Recall from chapter 1 that we began with the discrete form and derived the continuous form. Here we just do the reverse.

The One-Dimensional Map

The one-dimensional map (one-dimensional because only one dynamic state variable is under consideration) is a convenient modeling technique, especially because of its obvious graphical interpretation: it is possible to rapidly gain an idea of the dynamic behavior of a model simply by glancing at a graph. A one-dimensional map applies to those systems that can be represented as the projection of a variable from one time unit to the next. First construct a graph of the population density in year $t + 1$ versus the population density in year t . Suppose, for example, that the population density beginning in year 1997 is 10 and in subsequent years it is 20, then 40, then 80, then 100, then 110. That is, $N_{1997} = 10$, $N_{1998} = 20$, $N_{1999} = 40$, etc. To graph the numbers in the style of a one-dimensional map we first graph 10 on the abscissa and 40 on the ordinate, then 40 on the abscissa and 80 on the ordinate, then 80 on the abscissa and 100 on the ordinate. In doing so we are essentially making a graphic form of the number series 20, 40, 80, 100, 110. We know that the number 20 projects into 40, and drawing a vertical arrow from 20 on the abscissa to the intersection of a horizontal arrow from the value of 40 on the ordinate is simply a graphic statement of this fact (that 20 projects into 40). We now wish to project from 40, in which case we simply draw a similar arrow from 40 on the abscissa to the point where it intersects the value of 80 on the ordinate. The first projection (from 20) yielded 40, and we sought to initiate the second projection from this value of 40. This is a general rule. The next projection always begins where the previous projection left off. How can we know where that initiation is? We can obviously simply search for the ordinate value on the abscissa (i.e., after the first projection from 20 to 40, we search for 40 on the abscissa so as to make the second projection). But that search is made graphically much simpler if we draw a reference line beginning at zero for both abscissa and ordinate and rising at a 45 degree angle to the axes. This enables us to take the original projection and reflect it back to the 45 degree line. It is simply a graphic technique for locating the projected value on the abscissa so that it can be projected into the next time period. This whole example is illustrated in figure 4.16.

This simple example can now be generalized. Instead of using specific numbers, we may write a general rule of projection. For example, N at time t will become $N + 5$ at time $t + 1$ ($N_{t+1} = N_t + 5$), or N next year will be twice the value of N this year ($N_{t+1} = 2N_t$), or N next year depends on the value of N this year, that is, N next year is a function of N this year ($N_{t+1} = f(N_t)$). Although it is frequently possible to state the exact relationship between N this year and N next year, in the absence of that knowledge it is also useful simply to be able to draw the general shape of f , which is frequently pos-

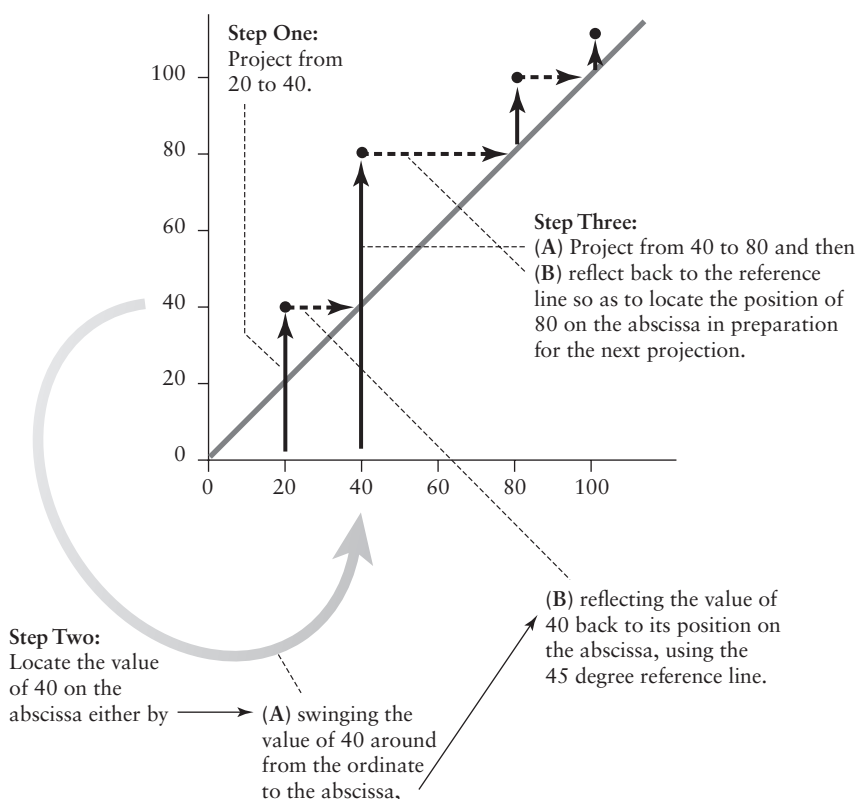


FIGURE 4.16. Step-by-step illustration of the process of stair-stepping using numerical values for a one-dimensional map.

sible only from qualitative knowledge of how the system behaves. But the stair-stepping procedure is still the same. Instead of projecting from 20 to 40 (as in the above example), project from an arbitrary starting point on the abscissa to the graph of the function. Then locate that projected value on the abscissa by reflecting it back to the 45 degree reference line and project it to the graph of the function again. This process is illustrated in figure 4.17.

The general rule, which is then repeatedly iterated, is project to the function, reflect to the reference line, project to the function, reflect to the reference line, and so on. After a short practice session, the general qualitative dynamics of almost any one-dimensional map can be rapidly visualized with a simple glance at the graph.

In figure 4.18, equation 1 is graphed along with the classical stair-stepping technique that can be used to quickly visualize the dynamics of the system. Where the graph of the equation crosses the 45 degree line, an equilibrium point exists. For equation 1 (the exponential equation), that equilibrium is at zero. If $\lambda > 1.0$, the particular nature of that equilibrium is unstable, because any value of N close to the point (i.e., the equilibrium $N = 0$) will deviate away from it. If the point were set at exactly $N = 0$, a glance at equation 1

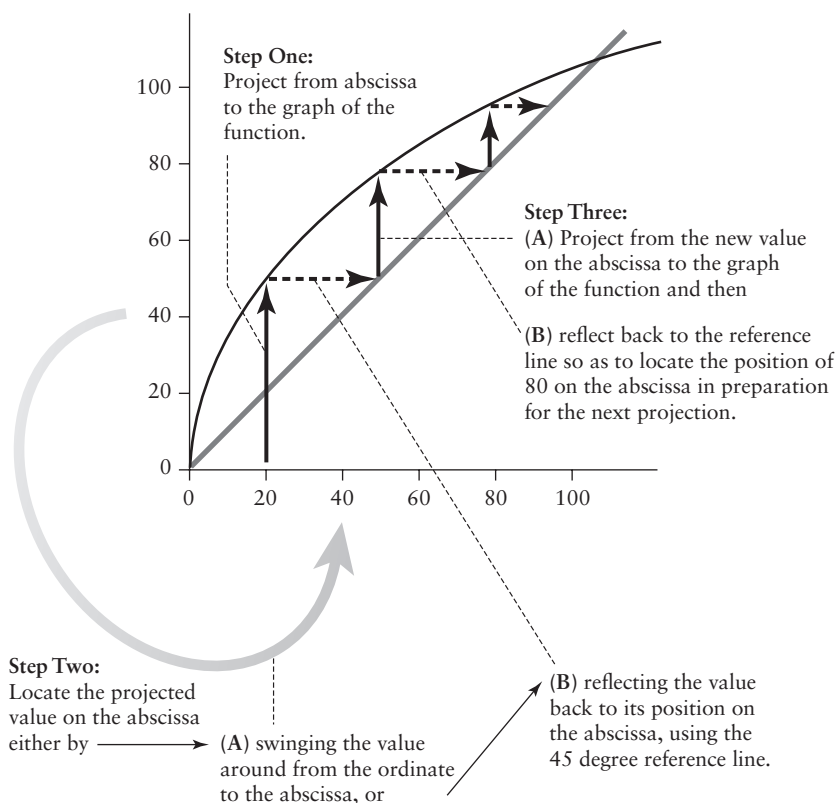


FIGURE 4.17. Step-by-step projection using a function rather than numerical values.

shows that it would stay there forever. But the slightest increase from zero (e.g., $N = 0.0000001$) means that the population will grow and thus deviate from the equilibrium point.

Now suppose that at each time interval a constant number of individuals migrates into the population. Suppose that the number is m , thus transforming equation 1 into

$$N_{t+1} = \lambda N_t + m. \quad (3)$$

A graph of equation 3 is presented in figure 4.19 (assuming that $\lambda < 1.0$). Once again the point at which the graph of the equation crosses the 45 degree line is an equilibrium point (setting N at exactly that point, which in this case is $m/(1 - \lambda)$, results in the same value of N for every future time period). This time, however, the equilibrium is a stable one, as illustrated in figure 4.19. Whatever the initial population size, the tendency will be to return to the value of $m/(1 - \lambda)$, the equilibrium state, which is thus an attractor.

Now assume that, instead of there being regular immigrants into the population, a predator population exists in the habitat, and that predator

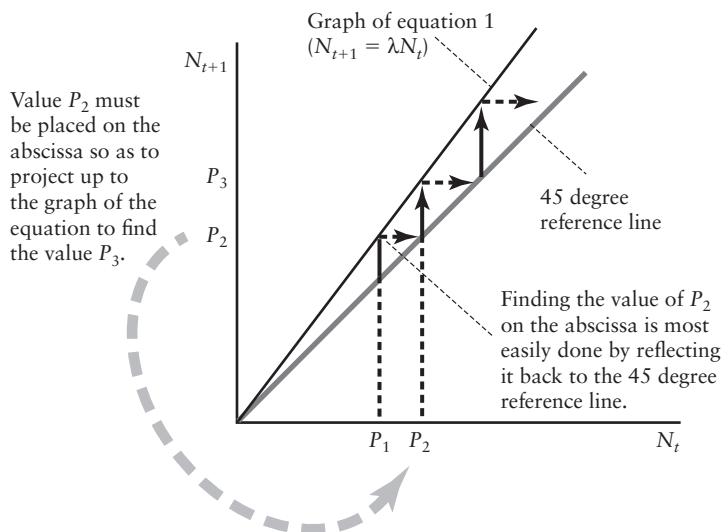


FIGURE 4.18. Exponential equation presented as a one-dimensional graph. The qualitative dynamics of such a system are easily visualized with the stair-stepping technique. Beginning at point P_1 , go up to the graph of the equation to reach P_2 on the ordinate. The ordinate value P_2 must then be positioned on the abscissa, which is most easily done by reflecting it to the 45 degree reference line (dashed arrow), which indicates its position on the abscissa. From P_2 on the abscissa, go up to P_3 and repeat the process (see figures 4.16 and 4.17).

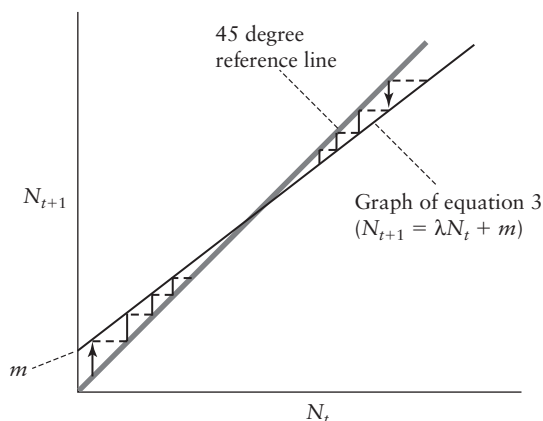


FIGURE 4.19. Graph of equation 3, illustrating a point attractor. The stair-stepping technique is the same as in figures 4.17 and 4.18. Any initiating point either above or below the attractor (where the graph of the function crosses the 45 degree reference line) eventually approaches that point. It is thus a point attractor, because any deviation from it will automatically revert to it.

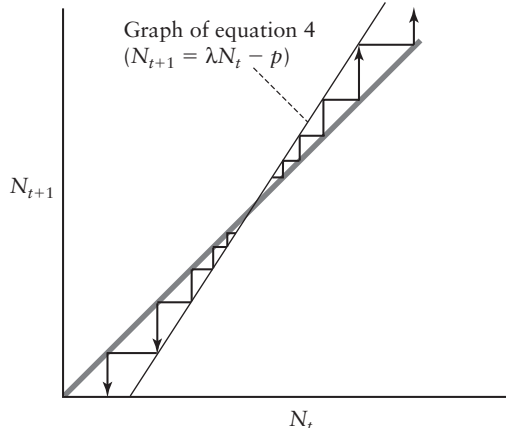


FIGURE 4.20. Graph of equation 4, illustrating an unstable equilibrium. The stair-stepping technique is the same as in figures 4.17, 4.18, and 4.19. Any point deviating only slightly from the equilibrium will continue deviating. It is thus a point repeller, because any deviation from it will continue deviating (it “repels” all values).

population does not change as a function of the prey density. Thus a constant number, p , of individual prey organisms will be taken out of the population in each time unit, and the appropriate equation is

$$N_{t+1} = \lambda N_t - p, \quad (4)$$

which is graphed in figure 4.20 (with the assumption that $\lambda > 1.0$). Note that the equilibrium point is $p/(\lambda - 1)$, and it is an unstable one, making the point a repeller, just as the zero equilibrium point was unstable for the original exponential equation (equation 1).

These two simple modifications to the basic exponential equation are both linear. However, most ecological processes of interest are known to be nonlinear, so it makes sense to modify equation 1 in a nonlinear fashion, too. The most elementary nonlinearity would be to assume that the parameter that multiplies the variable of interest (i.e., λ in equations 1, 3, and 4) is itself a function of the variable. If we assume that the parameter λ is a decreasing function of N (i.e., that the growth of the population depends on its density—recall density dependence from chapter 1) and furthermore that the exact function is $\lambda - \lambda N$ (i.e., the λ in equations 1, 3, and 4 becomes $\lambda - \lambda N$), the exponential equation (equation 1) becomes

$$N_{t+1} = \lambda N_t (1 - N_t). \quad (5)$$

In figure 4.21, equation 5 is graphed for two different values of λ . From the simple stair-stepping graphic technique it is obvious that both cases pictured are oscillatory. That is, at successive intervals the population alternately increases and decreases, as illustrated in the diagrams beneath the stair-stepped graphs. The difference between figure 4.21A and 4.21B is the difference between an oscillatory attractor (figure 4.21A) and an oscillatory repeller (figure 4.21B).

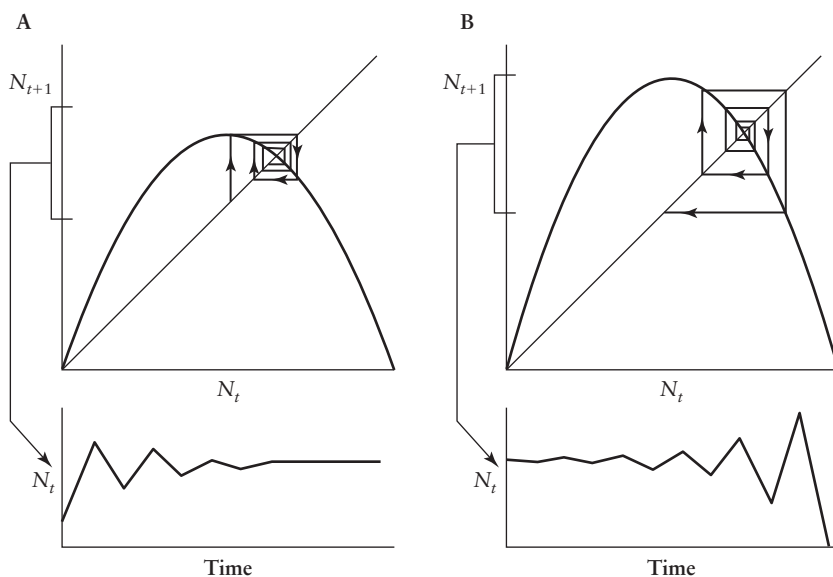


FIGURE 4.21. Graphs of equation 5. (A) Oscillatory attractor. (B) Oscillatory repeller. In both A and B the graph below the main graph illustrates the behavior of the variable through time.

EXERCISES

4.13 Graph N_{t+1} versus N_t for the logistic map for $\lambda = 3.4, 3.5, 3.56$, and 3.6 . Print out your graphs and make a pencil-and-paper stair-step diagram illustrating the dynamics of each graph. Compare these dynamics to those for the equivalent values of λ from exercises 4.10, 4.11, and 4.12.

4.14 The Ricker map is given generally as

$$N_{t+1} = \lambda N_t e^{1-bN_t}.$$

Plot N_{t+1} versus N_t for $\lambda = 6, 5, 4$, and 3 with corresponding $b = 2, 3, 4$, and 5 . Print out your graphs and make a pencil-and-paper stair-step diagram illustrating the dynamics of each graph.

4.15 Set up an Excel sheet to generate a logistic map with $\lambda = 4$. The first step should look like this:

	A
1	N
2	0.1000
3	=4*A2*(1-A2)
4	0.9216
5	0.2890
6	0.8219
7	0.5854
8	0.9708
9	0.1133
10	0.4020
11	0.9616
12	0.1478

Now set up a second column that simply reproduces the first one. Your sheet should look like this:

	A	B
1	N	N'
2	0.1000	0.1000
3	0.3600	=4*B2*(1-B2)
4	0.9216	0.9216
5	0.2890	0.2890
6	0.8219	0.8219
7	0.5854	0.5854
8	0.9708	0.9708
9	0.1133	0.1133
10	0.4020	0.4020
11	0.9616	0.9616
12	0.1478	0.1478
13	0.5039	0.5039
14	0.9999	0.9999

Now plot column A (labeled N) against N', which, of course, will yield a graph of all the points on the 45 degree line. Now take the sixth entry, which is equal to 0.5854, and retype it into the sixth position in column B (labeled N') (i.e., in place of the formula, directly type 0.5854). What happens to the graph? Why?

- 4.16** Modify the logistic map by subtracting a constant number (recall exercise 4.9), as might be done by a naïve manager. Let $\lambda = 3.5$, and subtract 0.25 individuals at each time step. Create a plot of N_{t+1} versus N_t and project the population 50 time units. Repeat for $\lambda = 4$, then for $\lambda = 4.55$.
- 4.17** For the modified logistic map of exercise 4.16, set $\lambda = 4.556$ and iterate over 200 time periods, beginning with an initial population of 0.3 and graph the time series (be sure to force your graph from 0 to 1).

Figures 4.19–4.21 summarize the classical notions of equilibrium and stability in one-dimensional maps. A single point is the equilibrium point, and it may be an attractor (stable) or a repeller (unstable), oscillatory or non-oscillatory. Because we are now dealing with discrete space rather than continuous space, the eigenvalue rules as elucidated in the previous section do not directly apply. The eigenvalue here is the slope of the function as it crosses the 45 degree line, with dynamics as summarized in figure 4.22, where it should be evident that an eigenvalue greater or less than 1 or -1 stipulates the qualitative nature of the equilibrium point. If the eigenvalue is >1 the system is a point repeller. If the eigenvalue is <1 but >0 the system is a point attractor. If the eigenvalue is <0 but >-1 the system is an oscillatory attractor. If the eigenvalue is <-1 the system is an oscillatory repeller.

If the system generates a strange attractor (or, for that matter, a permanent cycle), ideas of point attractors and repellers are useless, despite the fact that point repellers are always contained within strange attractors. For example, in figure 4.23 three cases are illustrated in which the critical equilibrium point (where the graph of the equation crosses the 45 degree line) is a repeller. However, knowing that the equilibrium point is a repeller provides us with scant information on what is interesting about the behavior of the system. Indeed, what is important in this case is the distinction between A and B on the one

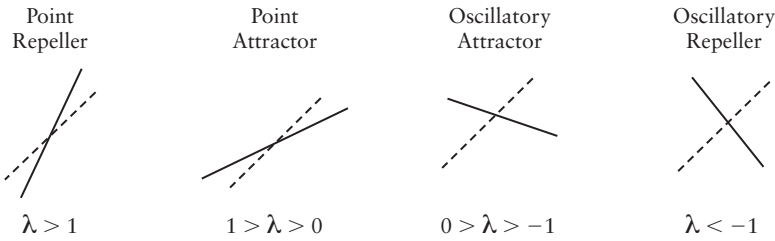


FIGURE 4.22. Eigenvalue values for the various forms of stability in a one-dimensional map. The dotted line is the 45 degree reference line, and the solid line is the function (illustrating that part of the function near to its crossing with the 45 degree line).

hand and C on the other. Cases A and B will persist indefinitely (i.e., are sustainable), whereas in case C the variable is extinguished from the system (the population goes locally extinct). If this were, for example, the population of an introduced natural enemy in an agroecosystem, we would care little that the population is theoretically unstable (i.e., its equilibrium point is unstable). Rather we would be concerned with whether the new natural enemy would persist in the environment, that is, whether we were dealing with, on the one hand, the situation in figure 4.23A or B or, on the other hand, with the situation in figure 4.23C. Our interest here would be not in the equilibrium point itself but in the limits, or boundaries, of the system. These boundaries are illustrated by dashed lines intersecting the two axes in figure 4.23A and B.

A further word is in order regarding the difference between the patterns in figure 4.23A and B. Figure 4.23A is classically known as an n -point cycle (the particular value of n in figure 4.23A is 2, because there are two actual values of N that repeat themselves forever, as indicated by the dashed lines crossing the axes). It is the one-dimensional equivalent of the classical limit cycle

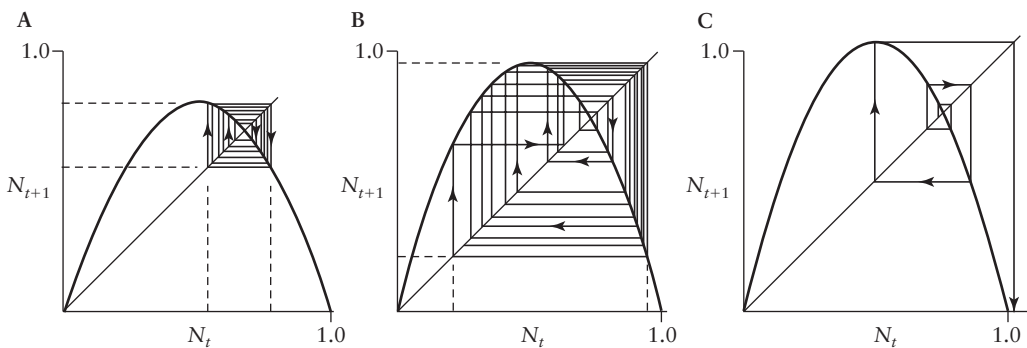


FIGURE 4.23. Graphs of equation 5. (A) A two-point periodic attractor ($\lambda = 3.1$). (B) A strange attractor ($\lambda = 3.8$). (C) An oscillatory repeller, leading to extinction of the population ($\lambda = 4.2$). In cases A and B the population has a repeller (is unstable) where the function graph crosses the 45 degree line, but in both cases the repeller is constrained by dynamic boundaries. Cases A and B are thus referred to as regionally stable.

in the context of differential equations. Figure 4.23B represents an example of chaos, better termed a strange attractor. The equilibrium point is unstable, yet there is no single n -point cycle, and trajectories generally move in a totally unpredictable direction, giving rise to the popular appellation *chaos*. An enormous literature is devoted to the analysis and significance of the difference between the pattern in figure 4.23A and that in 4.23B (e.g., Ellner and Turchin 1995; Hastings et al. 1993). To some extent that literature has been misdirected. The key question seems to have become whether ecological systems are chaotic. But to understand ecological systems it is not clear how knowing whether a system is chaotic or has a 50-point cycle will make much of a difference! True, a chaotic system is in theory unpredictable, but in practical fact an n -point cycle is just about as unpredictable if n is relatively large. On the other hand, in both figure 4.23A and figure 4.23B there are clear structures to the trajectories. Both are fundamentally oscillatory, even though the peaks of the oscillations do not repeat themselves exactly each year in the case of 4.23B and, most importantly, both have limits that they will never transcend (in a strictly deterministic world).

These limits are essentially identical to what Lewontin (1969) has referred to as dynamic boundedness. They define a section of the state space into which all nearby trajectories will eventually enter but that no trajectory can ever exit. Because all nearby trajectories must enter this space, the space itself is called an attractor, even though the equilibrium point within that space is unstable. Whether an attractor is strange or periodic will not be an important focus of the rest of this chapter. The significant practical feature for understanding ecosystem dynamics is the location of the boundaries and the qualitative structure of the dynamics, for both periodic and strange attractors (and repellers). Thus the important question to be asked of an unstable point is whether the nonlinearities of the system create boundaries around that point, thus making it either a periodic or a strange attractor. If not, the system will extinguish itself.

Stability and Equilibrium in the Logistic Map

The logistic map (as equation 5 is usually called), can be used to illustrate these and other simple ideas in a straightforward manner. The equilibrium point is

$$N^* = (\lambda - 1)/\lambda$$

(there is another, trivial, equilibrium point at $N^* = 0$). This means that λ must be greater than 1.0 to have a positive equilibrium point, and the equilibrium will be stable and nonoscillatory whenever the derivative of the function, evaluated at the equilibrium point, is greater than 0 (these conditions, and the ones that follow, should be clear after a detailed examination of the graph of equation 5 in figure 4.21; note especially that here the parameter λ is not equal to the eigenvalue, as it was in the case of the exponential equation). That is, differentiating equation 5, we obtain

$$(dN_{t+1}/dN_t) = \lambda - 2\lambda N_t,$$

which, when evaluated at the equilibrium point (i.e., substitute $N_t = (\lambda - 1)/\lambda$), is

$$(dN_{t+1}/dN_t) = \lambda - 2\lambda[(\lambda - 1)/\lambda] = \lambda - 2\lambda + 2 = 2 - \lambda,$$

in which case we see that as long as $\lambda < 2$ the equilibrium point will be a simple nonoscillatory attractor (note that the case of an unstable nonoscillatory point, as shown in figure 4.20, is not possible with this model). When $\lambda > 2$, the system will be oscillatory. It will have an attractor if the derivative of the function evaluated at the equilibrium point is greater than -1 . Thus

$$2 - \lambda > -1$$

or

$$2 < \lambda < 3$$

indicates an oscillatory attractor, and

$$2 - \lambda < -1$$

or

$$\lambda > 3$$

indicates an oscillatory repeller. Note that the value of the eigenvalue is $2 - \lambda$, making these observations consistent with the eigenvalue conditions of figure 4.22.

The existence of an oscillatory repeller when $\lambda > 3$ leads to the further question of how to distinguish between persistence (cases A and B of figure 4.23) versus extinction (case C of figure 4.23). Extinction will occur when the projection from the maximum value of the map falls on the x -axis at a point greater than the intersection of the function (see figure 4.24).

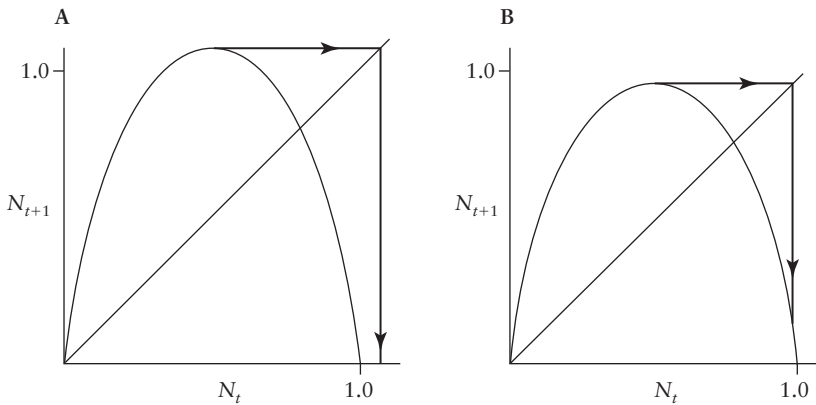


FIGURE 4.24. The difference between stability and instability in the regional sense. (A) Regionally unstable, the population will go extinct. (B) Regionally stable, the population will persist (albeit in a chaotic state). In both cases, the equilibrium point is unstable in the neighborhood sense.

From an examination of the logistic map (equation 5) we see that the function intersects the abscissa at 0 and 1. Thus extinction of the system will occur when the projection from the peak of the map touches the x -axis at a point greater than 1 (see figure 4.24A). The peak occurs at $N_t = 0.5$, so its projection (from equation 2) will be

$$N_{t+1} = \lambda(0.5)(1 - 0.5) = 0.25\lambda,$$

and the condition for extinction is thus

$$0.25\lambda > 1,$$

$$\lambda > 4.$$

When we combine this information with the earlier observation that the system will have an oscillatory (either periodic or strange) attractor whenever

$$3 < \lambda < 4,$$

we can say that the system will be sustainable as long as λ is between 1 and 4, even though the classic conditions for stability fail for $\lambda > 3$. Although the specific development in this chapter is associated with the density of a single population, the same dynamic rules apply if the state variable is some other interesting variable. For example, N might be the yearly production of manure from a dairy farm or the soil organic matter in a forest system or so forth. If we presume that equation 5 represents the system, we can unambiguously define sustainability as $1 < \lambda < 4$. The trick, of course, is that equation 5 is normally too simple to accurately represent anything as complicated as organic matter or manure (or even population density), and we use it here for didactic purposes only.

The upper and lower boundaries of the system are easily calculated. The upper limit is simply the peak of the function

$$N_{t+1} = \lambda(0.5)(1 - 0.5) = 0.25\lambda,$$

and its projection,

$$N_{t+1} = \lambda(0.25\lambda)(1 - 0.25\lambda) = 0.25\lambda^2 - 0.0625\lambda^3,$$

is the lower limit. Again, depending on the context, such boundaries may be of tremendous interest. For example, if N is the population density of a pest insect and the damage threshold is known (say it is D), the population will never be a pest if 0.25λ (the upper threshold) is less than D . Thus $\lambda < 4D$ exactly stipulates the conditions under which this population will be an occasional pest.

Basins of Attraction in the Logistic Map

For most simple models of ecological processes it has been possible to simply analyze the equilibrium point(s) and leave it at that. Most ecologists now admit that more complicated models are necessary to reflect even the simplest ecological phenomena. With even slightly more complex models we face a sit-

uation in which alternative equilibria exist in the same model. For example, if we combine the ecological principle that led to equation 4 with that which led to equation 5, we obtain

$$N_{t+1} = \lambda N_t(1 - N_t) - p \text{ for } N_t > 0, \quad (6a)$$

where λ is again the rate of population increase and p is the number of individuals removed from the population during each time unit by a constant predator. To make equation 6a relevant to ecological processes, we restrict its application to $N_t > 0$ and add the equation

$$N_{t+1} = 0 \text{ for } N < 0. \quad (6b)$$

This condition simply acknowledges that there can be no values of N less than zero (assuming the running example of N signifying the population density of an insect pest; other variables may take on negative values, in which case the special condition for $N < 0$ would not be necessary). Equation 6 is graphed in figure 4.25. There are three equilibrium points, given as

$$N^* = 0 \text{ (from equation 6b),} \quad (7a)$$

$$N^* = [(\lambda - 1)/2\lambda] + \{[(\lambda - 1)/2\lambda]^2 - p/\lambda\}^{1/2}, \quad (7b)$$

and

$$N^* = [(\lambda - 1)/2\lambda] - \{[(\lambda - 1)/2\lambda]^2 - p/\lambda\}^{1/2}. \quad (7c)$$

Inspecting figure 4.25, we see that the central equilibrium point is a repeller and the lower one (at $N = 0$) is an attractor, as is the upper one. Although

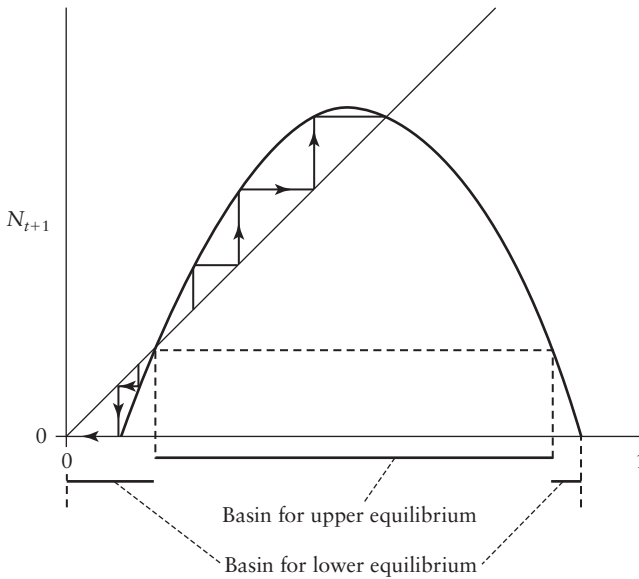


FIGURE 4.25. Graph of equation 6, illustrating the two basins of attraction for the two point attractors.

knowing the locations of the equilibrium points is important, another feature of figure 4.25 is important to understand the population dynamics. Any value of N near to but greater than the repeller will eventually approach the upper attractor, whereas any point less than the repeller will eventually approach the lower attractor. The repeller thus separates the state space (all possible values of N) into those values that approach the upper attractor and those values that approach the lower attractor (this is only approximately true, as discussed in the next paragraph). The unstable point in this context is referred to as a separatrix, and the collection of points on either side of it is a basin of attraction. The basin of attraction refers to the section of the relevant state space in which all trajectories approach a given attractor. This question of which of the initial values will eventually reside in particular locations in the state space may turn out to be far more important for analyzing ecosystems than is the traditional question of the exact location of the equilibrium point and whether it is stable (see, for example, Scheffer et al. 2009).

This issue is a bit more complicated in the case of the logistic map when the value of λ is very large. Very high values of N , because of the strong density dependence of the logistic model, will be projected to values of N less than the separatrix. Thus there is a section of the lower equilibrium's basin of attraction that exists at very high values of N , in addition to the obvious one that exists at lower values of N , as discussed later.

EXERCISES

- 4.18** Generate a time series using the formula for density-dependent population growth of Bleasdale and Nelder (1960) and Hassell (1975),

$$N_{t+1} = \frac{\lambda N_t}{1 + N_t^b},$$

setting $\lambda = 5$ and $b = 4$ and reiterating for 50 time units. Create a graph of the function from the equation, and experiment with other values of λ and b (in the spreadsheet, fix λ and b separately, then generate both the function graph and the time series).

- 4.19** Repeat exercise 4.18 but with $\lambda = 1$ and $b = 1$. Experiment with a variety of values of b . What do you conclude about the qualitative behavior of the model with respect to variation in the parameter b (try $b = 1, 1.5$, and 0.5 , for example)?
-

Structural Stability

A notion of stability totally distinct from that discussed so far may arise when parameters undergo change. That is, in all the above examples, the state variable (X_t or Y_t or N_t), the one that is dynamic, which is to say the one that varies through time, is clearly distinguished from the parameters, which do not vary through time. For example, in equation 6a, N_t is the state variable, while λ and p are parameters. For purposes of analysis we presume that N_t varies through time, while λ and p do not.

A different sort of analysis emerges when we ask what happens when λ and p themselves vary within a distinct time frame. For example, N_t may vary in ecological time while λ changes slowly as evolution forces change in its value, and p may change as the resident predator population slowly increases or decreases. It is of great interest to examine what happens to the general results as the parameters change and what this has to do with stability. Thus “state space” and “parameter space” are quite distinct concepts. State space is represented as a graph of the potential value(s) of the state variable(s) at a given set of parameter values, whereas parameter space is represented as a graph of the potential values of all the parameters in the model.

Consider the case of a nonreproductive population that receives migrants in a density-dependent fashion. That is, suppose that the rate of migration into the population is f , but f itself is a negative function of population density (the migrants have the ability to sense when the population is overcrowded, for example, and tend to avoid an overcrowded situation). This circumstance could be modeled with the simple equation

$$N_{t+1} = \lambda(1 - N_t). \tag{8}$$

As illustrated in figure 4.26, if the value of λ is greater than 1.0, the equilibrium point is oscillatory and a repeller (figure 4.26A). If the value of λ is less than 1.0, the equilibrium point is oscillatory and an attractor (figure 4.26C). The question then arises, What if the value of λ is exactly 1.0? Such a situation presents precisely the behavior one would expect mathematically: oscillatory and neither an attractor nor a repeller (figure 4.26B).

Although it may not seem particularly important that the population permanently oscillates between two particular values, the form of oscillation is particularly unusual. At every second time projection, the population will return to exactly what it had been before, no matter where it started. For example, if we begin with $N = 0.3$, the next value will be 0.7 (see equation 8) and the next value 0.3 again, whereas if we begin with 0.2, the next value will be 0.8 and the following one 0.2 again. That is, the population will oscillate with a cycle that is two time periods in length, but the exact values of the cycle will depend

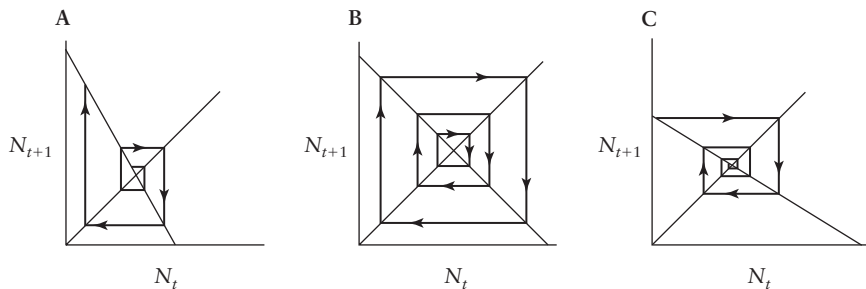


FIGURE 4.26. Illustration of a structurally unstable parameter configuration for equation 8. (A) A point repeller resulting from a slightly larger value of λ than in B. (B) A neutrally stable situation in which the initiation point is forever repeated every other time unit. (C) A point attractor resulting from a slightly smaller value of λ .

on the starting point. Although this situation is thought to be uninteresting in a biological sense, it is actually rather important mathematically and is ultimately a key point for conceptualizing the idea of structural stability. The behavior is illustrated in figure 4.26B, in which we see that the oscillations are neither stable nor unstable. This state is dependent on the assumption that $\lambda = 1.0$. If $\lambda = 0.9999$, say, the population will no longer continue returning to the point at which it started but rather will slowly converge on a value of about 0.67 (figure 4.26C; actually the value is about 0.66665555). That is, the qualitative behavior of the system changes dramatically when the value of λ is changed only slightly, from a system that is dependent on the starting point, ever cycling back to the same point, to a system that converges on a single equilibrium point—an attractor. Similarly, if $\lambda = 1.00001$, the system will slowly oscillate away from the point 0.67, an oscillatory repeller (figure 4.26A). So the value of $\lambda = 1.0$ is a kind of break point for the parameter λ . When λ is either greater than or less than 1.0 the system has qualitatively distinct behavior. Points such as $\lambda = 1$ in equation 8 are described as structurally unstable (or the model is structurally unstable at that point) because the slightest change in the parameter will yield a qualitatively distinct form of behavior for the system in general.

Points of structural instability, also called bifurcation points, often occur in ecological models, especially in discrete time, and they play a crucial part in analyzing the overall qualitative behavior of models. For another example, returning to the logistic equation (equation 5), three situations are illustrated in figure 4.27: $\lambda < 2.0$, $\lambda = 2.0$, and $\lambda > 2.0$. In the same sense as above, $\lambda = 2.0$ appears to be a structurally unstable situation. The smallest reduction from the value of 2.0 yields a population that asymptotically approaches a point attractor, whereas the smallest increase from the value of 2.0 yields a population that oscillates toward a point attractor. Thus the model is structurally unstable when $\lambda = 2.0$.

Another structurally unstable point is illustrated in figure 4.28. In this case the middle figure (figure 4.28B) is a graph of the logistic with $\lambda = 3.0$. If λ is decreased slightly, the figure in figure 4.28A emerges and the behavior of the

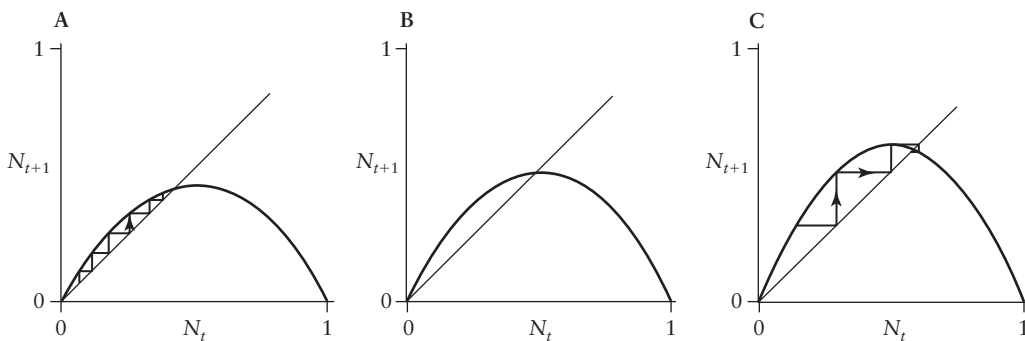


FIGURE 4.27. Graphs of equation 5 (the logistic map), illustrating the structurally unstable configuration obtained when $\lambda = 2.0$. (A) $\lambda < 2.0$, leading to a stable node (nonoscillatory point attractor). (B) $\lambda = 2.0$, the bifurcation point. (C) $\lambda > 2.0$, leading to a stable focus (oscillatory point attractor).

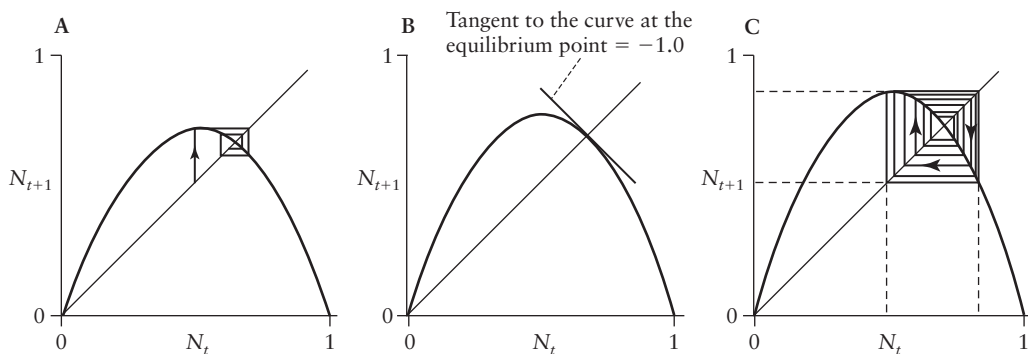


FIGURE 4.28. Graphs of equation 8 (the logistic map), illustrating the structurally unstable configuration obtained when $\lambda = 3.0$. (A) $2.0 < \lambda < 3.0$ (see figure 4.27C). (B) $\lambda = 3.0$. (C) $\lambda > 3.0$, leading to a two-point attractor.

system is damped oscillations to a point attractor. If λ is increased from 3.0, the figure in figure 4.28C emerges and the behavior of the system is permanent oscillations with a period of 2, which is to say that the population, no matter where it is initiated, eventually oscillates forever between two fixed values, indicated with dashed lines in the figure. Again, the model with $\lambda = 3.0$ is structurally unstable in the sense that $\lambda = 3.0$ is a bifurcation point, with qualitative changes in the behavior of the system emerging when the parameter is changed ever so slightly from this value. The system switches from a single stable point attractor (oscillatory) to a stable period-two cycle. This type of bifurcation is known as a period-doubling bifurcation (in the context of continuous systems, i.e., with differential equations, the same qualitative arrangement is known as a Hopf bifurcation).

A very different type of bifurcation may arise in more complicated models. Consider, for example, the case modeled above of a constant population of predators in a system (equations 6a and 6b). With the appropriate choice of parameters, the situation in figure 4.29 may arise. Once again, the center graph

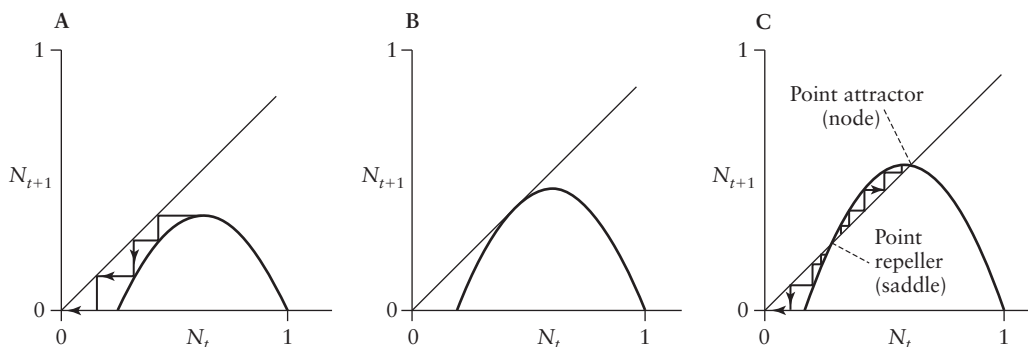


FIGURE 4.29. Illustration of a saddle-node bifurcation. (A) A single attractor at zero. (B) The bifurcation point. (C) After the bifurcation, there is a repeller (saddle) and an attractor (node), indicating that the bifurcation was of the saddle-node type.

(figure 4.29B) is a bifurcation point in that one small change in a parameter may create the situation in figure 4.29A, whereas a change in the other direction might result in the situation illustrated in figure 4.29C. Note that two new equilibrium points have been created (or destroyed) in this bifurcation, one an attractor (sometimes referred to as a node) and one a repeller (sometimes referred to as a saddle). This type of bifurcation is variously referred to as a saddle-node bifurcation or as a blue-sky bifurcation (because two equilibrium points appear “out of the blue”).

EXERCISES

4.20 Using the logistic map,

$$N_{t+1} = \lambda N_t(1 - N_t),$$

solve for the equilibrium value of N and plot the equilibrium value as a function of λ .

4.21 Using the modified logistic map with constant predation,

$$N_{t+1} = \lambda N_t(1 - N_t) - p,$$

solve for the equilibrium value of N and plot the equilibrium values (that’s plural) as a function of λ . Generate two time series with the model using $\lambda = 3$, $p = 0.3$, starting with $N_0 = 0.3$ and $N_0 = 0.2$. Generate two other time series with the model using $\lambda = 4.4$, $p = 0.3$, $N_0 = 0.11$, and $N_0 = 0.10$. Compare the time series to what you would have expected from the graph of the equilibrium values as a function of λ .

In recent years the saddle-node form of bifurcation has attracted a great deal of attention because of the fundamental idea of a regime shift (Scheffer 2009; Scheffer et al. 2009, 2012). As is evident in figure 4.29, alternate states exist for the system, and those alternate states may, in the real world, be alternate forms of an ecosystem, sometimes with important practical consequences. For example, recently it has been suggested that some of the world’s most common terrestrial formations are actually alternate modes of ecosystem organization, such that whether a system is a desert or a savannah may be a consequence of the point of initiation (figure 4.30). One of the implications is the possibility of “tipping points” that will rapidly shunt the system into one or another of the states, a so-called regime shift. Furthermore, there are reasons to expect a hysteresis, a range of some parameter value for which the system effectively gets stuck in one regime. So, for example, if climate change continues to produce drier conditions at the south end of the Sahara Desert, much of the savannah will convert to desert, yet if the climate were then to reverse and became more moist, the desert would persist.

The above two types of bifurcation (period-doubling and saddle-node) are both characteristic of changes in point attractors or repellers. In the case of the period-doubling bifurcation, the bifurcation itself shifts the model from a point attractor to an oscillatory attractor, but before the bifurcation, the equilibrium is a point attractor. Other types of bifurcation may involve strange

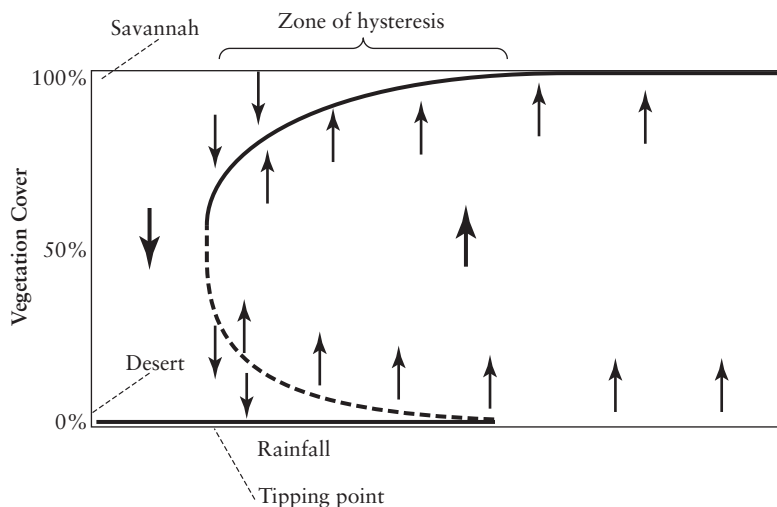


FIGURE 4.30. Example of a possible application of a saddle-node bifurcation (also more popularly known as a tipping point). Note that the vegetation cover is represented as a bold line, either a stable set of points (a node) with a solid bold line or an unstable set of points (a saddle) with a dashed bold line. As rainfall becomes less abundant, the vegetation cover slowly declines. However, there appears to be a tipping point, which is to say a point at which the vegetation cannot be sustained at all, and a desert results. This sort of arrangement also implies a zone of hysteresis; as the rainfall decreases, the system suddenly becomes a desert, but once it is in a desert, even though rainfall might increase again, the desert persists over the whole range of hysteresis.

attractors and tend to be far more complicated. Consider, for example, the constant predator model (equations 6a and 6b). In figure 4.31 this model is plotted first (figure 4.31A) with parameters such that there are two attractors, to the left a point attractor at zero and to the right a strange attractor, and the two are separated by a separatrix. A small change in parameter results in the condition illustrated in figure 4.31B. A further small change in parameter causes the strange attractor to collide with the basin of the point attractor, eliminating the strange attractor entirely (figure 31C). What used to be parts of the trajectory of the strange attractor are now simply trajectories within the basin of attraction of the point attractor at zero. This sort of bifurcation phenomenon is known as a basin boundary collision because the basin of one attractor (the lower point attractor at zero) collides with the dynamic boundary of the strange attractor.

These examples, two structural instabilities associated with point attractors (period-doubling and saddle-node bifurcations) and a structural instability associated with a strange attractor (basin boundary collision), are but a few of the possibilities. Other structural instabilities are possible with more complicated models but are beyond the scope of this text.

Structural stability, then, is a very different form of stability. It involves the structure of the model as a whole and its qualitative behavior.

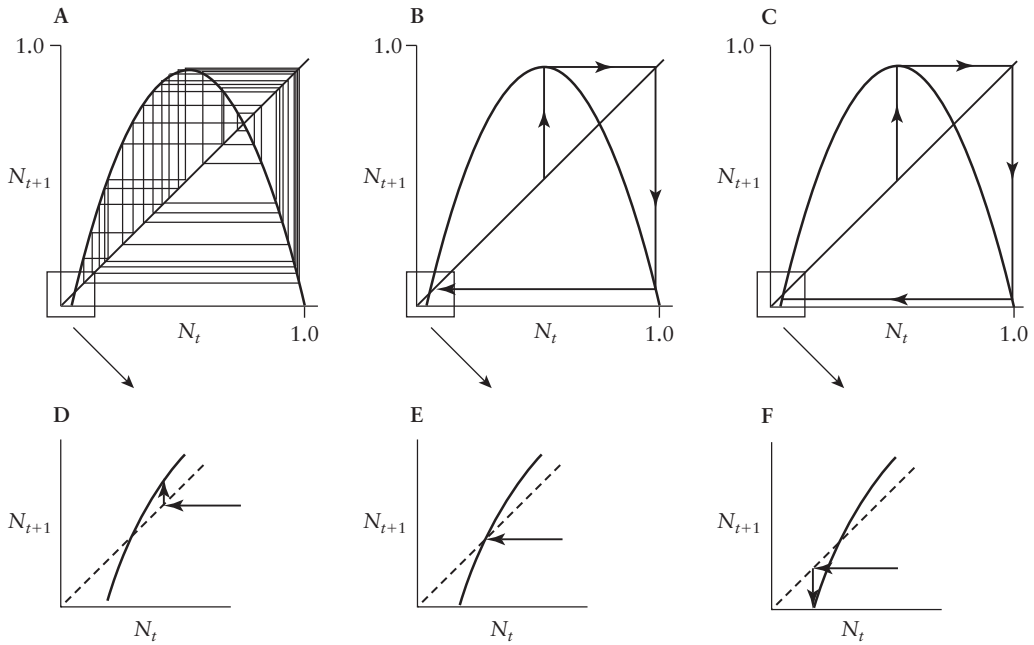


FIGURE 4.31. Illustration of a basin boundary collision. (A) Alternative attractors, one at zero and the other a strange attractor. (B) Bifurcation point. (C) After the basin boundary collides with the strange attractor, the trajectories that had been part of the attractor are now just part of the basin of attraction for the point attractor at zero. Panels D, E, and F are microscopic views of the area near zero, illustrating the positions of the trajectory emanating from the peak of the function.

Bifurcation Diagrams

Nonlinear models are frequently quite resistant to traditional analytical treatment. The one-dimensional maps (discrete time) and simple differential equations (continuous time) that have served to illustrate basic model behavior in this chapter are really only heuristic devices, enabling a partial understanding of the kinds of underlying structures that may give rise to complicated behaviors in more realistic models. For example, in more complicated predator–prey models, it is sometimes possible to construct an approximate one-dimensional map that captures the qualitative behavior of the more complicated model (e.g., Schaffer 1985). The basic structure of the one-dimensional model is then much easier to comprehend than the complicated model (a model of a model, so to speak).

One way of examining the general behavior of a complicated model when traditional analytical procedures are unavailable (i.e., because of the complexity of the model it is not possible to treat them analytically) is the bifurcation diagram. The various forms of bifurcation already described previously in this chapter (i.e., points of structural instability) are sometimes easily visualized in a “bifurcation diagram.” Consider the standard logistic map,

$$N_{t+1} = \lambda N_t(1 - N_t).$$

As shown earlier, this model will have a bifurcation point at $\lambda = 3$, with a simple equilibrium point when $\lambda < 3$ and a periodic solution when $\lambda > 3$. Specifically, the periodic solution is a “period-two” attractor, which is to say that the system oscillates between two points, exactly the same two points, forever. Consider, for example, the value of $\lambda = 3.2$. If we begin with

$$N_t = 0.8,$$

we can easily calculate

$$N_2 = 3.2N_1(1 - N_1) = 3.2(0.8)(1 - 0.8) = 3.2(0.8)(0.2) = 0.512.$$

Then we can substitute 0.512 to compute N_3 , as follows:

$$N_3 = 3.2(0.512)(1 - 0.512) = 3.2(0.512)(0.488) = 0.8,$$

which is the same number we originally started with. Thus we see that in this case the model will oscillate between 0.512 and 0.8 in perpetuity. This is a period-two (two values that are repeatedly visited) attractor.

In addition to the bifurcation point at $\lambda = 3$, without further proof, if λ becomes still larger, the period-two attractor converts into a period-four attractor, and if it becomes even larger, the period-four attractor converts into a period-eight attractor. (The interested reader can verify any of this with some simple experiments on a spreadsheet, following the approach of exercise 4.21.) In figure 4.32 this process is illustrated by way of a graph of N^*

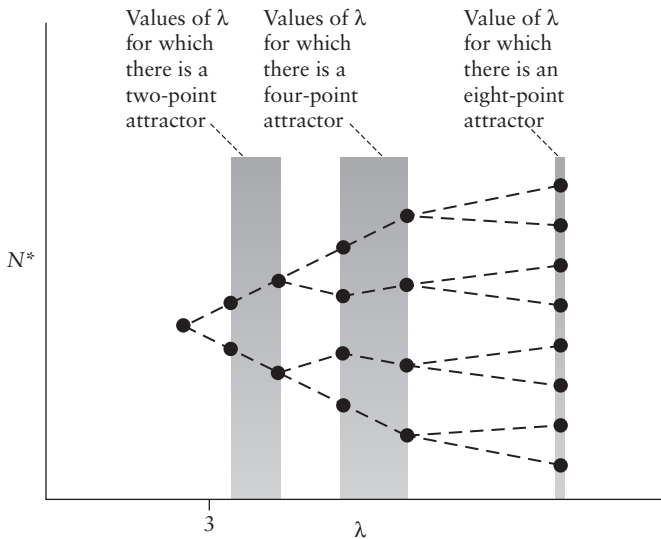


FIGURE 4.32. Illustration of the basic bifurcation process. Values of λ that result in various attractor types are illustrated. Dashed lines connect these points in what is likely to be the intermediate values giving rise to the different attractors. Note the bifurcating nature of the picture, giving rise to the appellation bifurcation diagram.

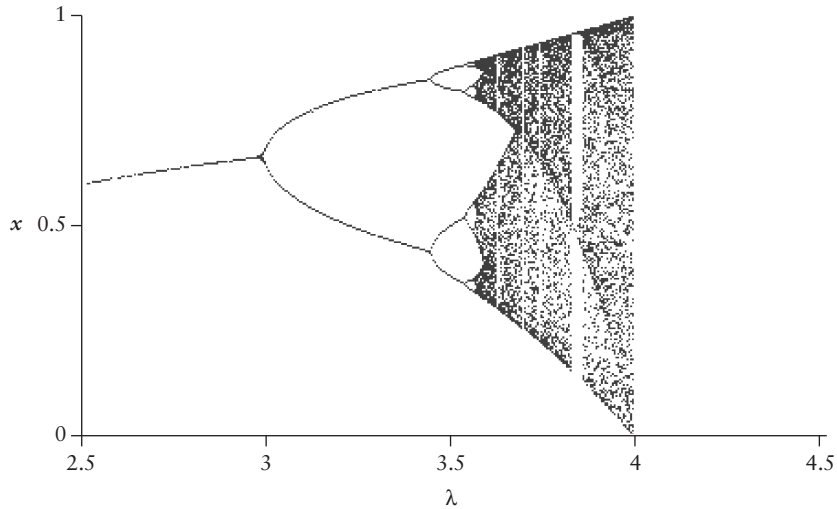


FIGURE 4.33. Bifurcation diagram for the logistic map. Note the similarity to figure 4.32. There is a clear period-doubling bifurcation at $\lambda = 3$. At $\lambda = 3.5$ another period doubling has already occurred and the system is in a four-point attractor. At λ slightly larger than 3.5 another doubling occurs and the period eight attractor is visible. Note also the period-three “window.” At $\lambda = 4$ there is another bifurcation, as described earlier in the text.

against λ , where N^* is either the single equilibrium point for a point attractor or one of the repeated points of the periodic attractor.

In figure 4.32 the points referred to above are plotted, but they are also connected with dashed lines. It is intuitively obvious that the dashed lines represent an approximation of what the intermediate values of N^* would be, and it is clear that the dashed lines bifurcate at critical points. The diagram in figure 4.32 is thus referred to as a bifurcation diagram, and is an important tool used to study complex models.

Rather than choosing particular values of λ and calculating the values of N^* , we can simply calculate N^* for all of the values of λ , incrementing by some small amount. If we do this for the logistic map, we obtain the graph presented in figure 4.33.

The various attractors described in figure 4.32 are clearly visible in figure 4.33. Also visible is the bifurcation event at $\lambda = 4$, fully explained in an earlier section. By examining such a bifurcation diagram it is frequently possible to gain an overall picture of how the model behaves. In this case there is a clear cascade of period-doubling events, from one to two to four to eight. In figure 4.34 a part of the bifurcation diagram is expanded ($\lambda = 3.5 - 3.7$).

Note that the period-doubling cascade is now visible for periods 4, 8, and 16. What happens is what one would expect, for the most part. The periods keep doubling, and the change necessary in λ to get to the next doubling keeps decreasing. Eventually there has been such a massive period doubling that a remarkable point is reached at which one can simultaneously get all

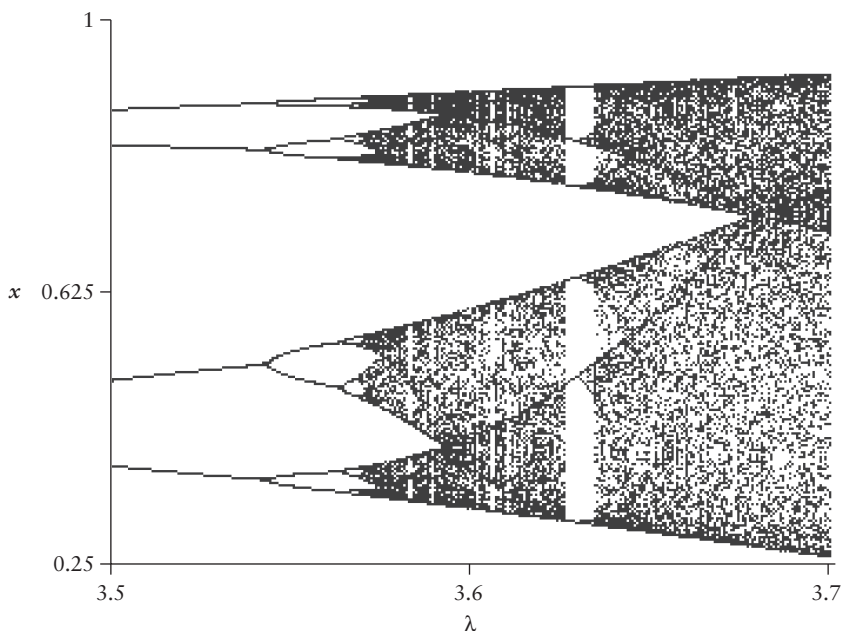


FIGURE 4.34. Close-up of part of the bifurcation diagram of figure 4.33.

possible periods as well as an uncountable number of aperiodic (i.e., never settling down to permanent values) attractors. This is the point at which the system is usually referred to as chaotic, and the manner in which chaos is approached here is referred to as the period-doubling route to chaos (there are other, qualitatively distinct, forms, but those are beyond the scope of this text). Just how complicated this behavior is can be appreciated by noting that there is a “window” in the original diagram (figure 4.33) in which there is an attractor of period 3. Where does a period-three attractor come from if the sequence goes $1, 2, 4, 8, 16, \dots$? Similarly, in figure 4.34 there is a window with a clear period-6 attractor. Where does that come from? Suffice it to say here that the explanation is mathematically complicated and beyond the intentions of this text. But biologically all it means is that the system is extremely unpredictable in the range of about $3.58 < \lambda < 4$, which is why the term *chaos* seems to be so popular.

As an example of the utility of bifurcation diagrams, consider the *Tribolium* model presented in chapter 2. Recall that the population was divided into larvae, pupae, and adults and the nonlinear projection matrix model was given as

$$\begin{vmatrix} L_{t+1} \\ P_{t+1} \\ A_{t+1} \end{vmatrix} = \begin{vmatrix} 0 & 0 & f_1(L_t A_t) \\ p_{lp} & 0 & 0 \\ 0 & f_2(A_t) & p_{aa} \end{vmatrix} \begin{vmatrix} L_t \\ P_t \\ A_t \end{vmatrix},$$

where L is the number of larvae, P is the number of pupae, and A is the number of adults. The functions f_1 and f_2 stipulate the nonlinear effect of cannibalism on the production of larvae by adults and the survival of pupae to adulthood, respectively. These functions are given as

$$f_1 = \frac{b}{e^{c_1 L_t + c_2 A_t}}$$

and

$$f_2 = \frac{b}{e^{c_3 A_t}},$$

which are intended to incorporate the biological fact of cannibalism. This model would in fact be quite daunting if one were to try to solve it analytically, and it certainly does not lend itself to any obvious intuitive or heuristic explanation. But if one generates a bifurcation diagram, as Costantino et al. (1997), did, the diagram as pictured in figure 4.35 is obtained.

This bifurcation diagram was obtained by performing a series of experiments to estimate the parameters of the model and then substituting those values into the model, fixing all parameters except c_3 . The parameter c_3 represents the rate of consumption of larvae per adult and is a parameter that Costantino and colleagues could experimentally manipulate in the laboratory. They next chose particular values of c_3 that represented various different dynamic situations, as indicated by the arrows on the top of the bifurcation diagram, and set up laboratory cultures corresponding to those particular values of c_3 . Their results are shown in figure 4.36.

The open circles are the experimental results, and the closed circles are the expected results based on the model. The six different graphs correspond to

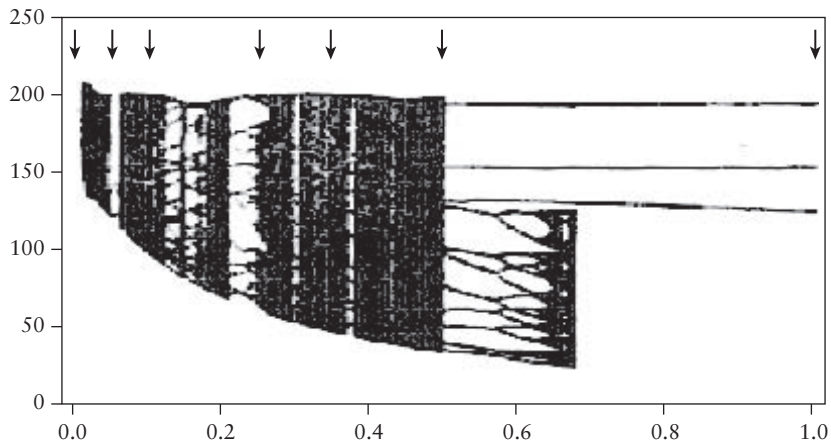


FIGURE 4.35. Bifurcation diagram of the *Tribolium* model. The bifurcation parameter is c_3 , consumption of pupae by adults, and the variable plotted is the total population size. Reprinted with permission from Costantino et al. (1997). © 1997 American Association for the Advancement of Science.

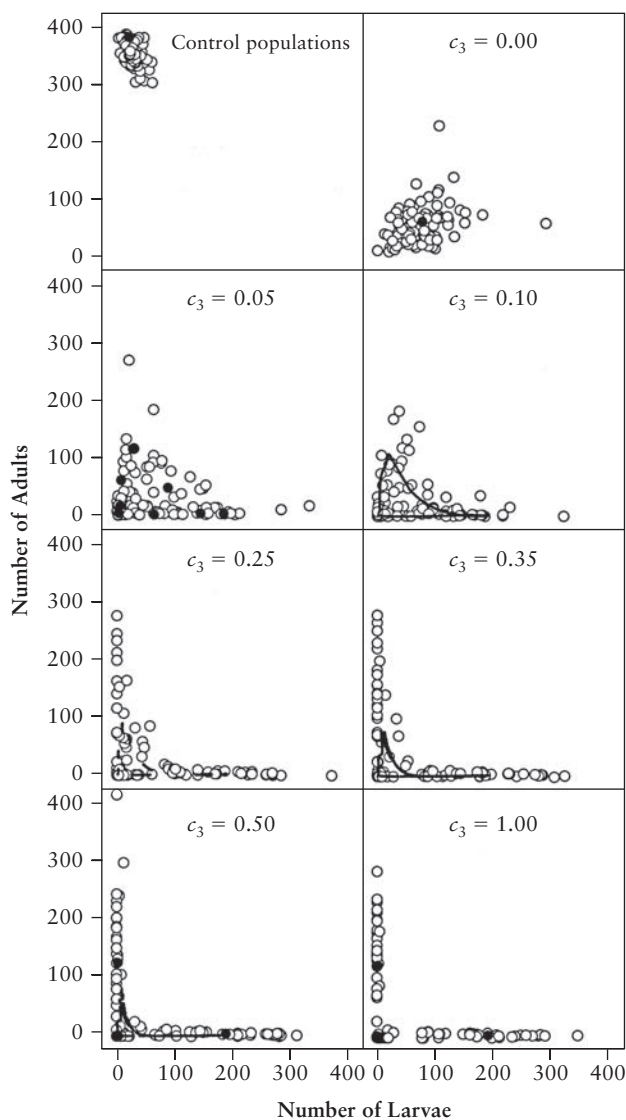


FIGURE 4.36. Experimental results of the *Tribolium* experiment of Costantino et al. (1997). The upper left-hand graph is the control; all others correspond to the values of c_3 indicated by the arrows in the upper part of figure 4.35. Open circles are the observed values; closed circles, lines, or loops are the expectation from the model. Reprinted with permission from Costantino et al. (1997). © 1997 American Association for the Advancement of Science.

the positions of the arrows in figure 4.35. There is a remarkable correspondence between what was expected and what is observed. What is even more remarkable is that the biologists on this team of investigators first thought that several of the theoretical outcomes would be impossible to achieve in the laboratory. But the data speak for themselves. Here the bifurcation diagram

was critical to the evaluation of the model to the point that predictions from the model came directly from the bifurcation diagram. The more traditional technique of comparing changes in the variables over time to the predictions over time generated by the model and by experiments could never have provided such a strong and elegant test of the model as did the analysis of the bifurcation diagram.

Concluding Remarks

In this chapter we have dealt with a panoply of different subjects related to the analysis of population models. These sorts of methods are currently the subject of intense investigation, and certainly this chapter will seem outdated within a few years. Nevertheless, these concepts are playing an increasingly important role in population modeling these days, and a minimal introduction, such as we have provided in this chapter, is essential background for understanding and using contemporary ecological approaches to population dynamics.

The primary focus in this chapter has been on the tools with which we can analyze models with complex behaviors, but the insights gained also lead to some important practical conclusions. We opened this chapter with notions of equilibrium and stability, long the bastions of creative thought in the search for a theory of ecosystems. The analyses in this chapter suggest that these notions may need to be abandoned, at least in terms of their classical meaning. For example, a dynamic system in chaotic behavior, such as that illustrated in figure 4.23B or figure 4.31A, is by classical standards unstable and in a strict mathematical sense unpredictable. Nevertheless, it has clear boundaries to its behavior and in another sense is quite stable—within the dynamic boundaries of its own region. Which sense is important to an ecosystem manager? Which sense is important in terms of understanding the ecosystem? Which sense is important in the context of natural selection?

These concepts may also suggest resolutions of various paradoxes of ecology. For example, the conundrum presented in the first paragraphs of this chapter, in which diversity is thought to generate stability yet some highly diverse systems are thought to be quite fragile, can be easily resolved. Perhaps the stability originally thought to result from the diversity actually refers to regional stability with a broad basin of attraction. The fragility that was actually observed when analyzing the question might then refer to the possibility that the basin itself could become smaller as diversity is reduced, increasing the likelihood that the basin could be traversed and the integrity of the system thus breached. Whether this is actually true of highly diverse systems is not the point here. Rather, these concepts suggest the way that some, perhaps many, natural historians have thought about this issue when they have pondered the relationship between diversity and stability.

In the sort of truly complex ecosystems likely to be encountered in the real world, the examples in this chapter will seem overly simplistic. Yet the Newtonian notion of point stability remains a stalwart of many thinkers in the

field of ecosystem dynamics, if only tacitly so. The simple one- and two-dimensional examples of this chapter are intended to introduce the notion of regional stability and the various complexities associated with it. Any real-world system will be multidimensional and ultimately must be represented in hyperspace. Figure 4.37 presents a “collapsed” hyperspace, a fictional two-dimensional representation of a multidimensional space. In figure 4.37A is a system with four unstable points, yet there are two attractors that contain two of the four unstable points. The attractors are regionally stable. According to classic definitions of stability, this would be a very unstable system indeed. Traditional analysis of neighborhood stability would determine that there are four equilibrium points, all of which are unstable. Yet almost anyone would agree that this system is more “stable” in some vague ecological sense than the alternative system illustrated in figure 4.37B. This system has two stable points (one at zero), yet intuitively most would regard it as less stable than the one in figure 4.37A. Clearly, the notion of regional stability more clearly encompasses what most workers in ecology would describe as stable, and a regional attractor is similarly closer to intuitive notions of equilibrium than are the single points of the neighborhood stability sense.

The notion of structural stability represents a totally different idea of stability than does either the neighborhood or the regional sense and in a variety of ways is probably similar to what many in the environmental movement really mean when they refer to a stable system—a system that shows particular characteristics and will continue to show those characteristics even if small changes in conditions occur in the environment. So, for example, when the local environment changes such that a crop pest develops a locally elevated population density, the natural enemies of the traditional agroecosystem may respond by exerting control over that temporarily elevated population. A change had occurred in a parameter (the local environment that resulted in

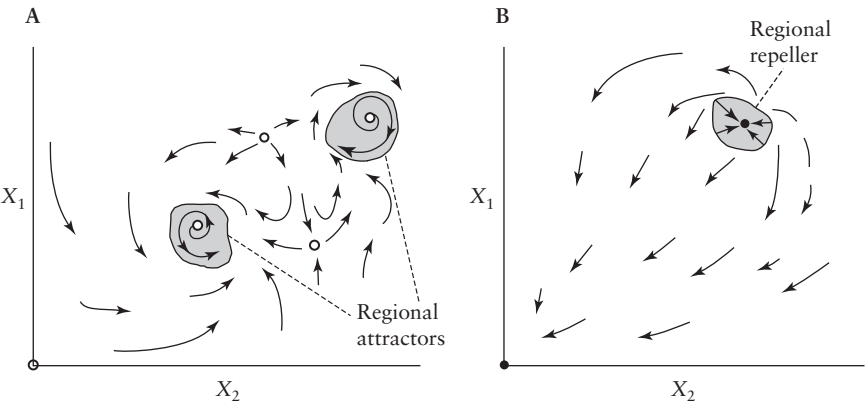


FIGURE 4.37. Theoretical situations in which (A) five-point repellers and two regional (strange) attractors coexist and (B) two-point attractors and a regional repeller coexist. The point attractors are illustrated by closed circles and the point repellers by open circles.

the elevated population density of the pest), but the system was structurally stable. No large change in its behavior resulted from this change in parameter.

Yet points of structural instability, especially bifurcation points, are sometimes exactly what we are looking for when we aim to understand or design ecosystems. Could it be, for example, that the conventional techniques for producing and processing tomatoes in California simply represent a syndrome of production, that another syndrome (the organic method, for example) might emerge if the parameters were changed somewhat, and that strategists aiming to convert to organic production might very well look for that break point, the bifurcation point that will carry the entire system into the organic mode? On the other hand, gradual changes in parameters may lead to a basin boundary collision in which a system originally held within a bounded attractor that represents a desired state of the ecosystem is engulfed in the basin of an attractor that includes undesired states. This phenomenon has been suggested as a possible mechanism for sudden extinction in natural populations (McCann and Yodzis 1994). This somewhat philosophical point will be left for the reader to ponder. Suffice it to say that the notion of structural stability is crucial in many ways to understand and/or design ecosystems.

Throughout this chapter we have minimized use of the word *chaos*, even though much of what is included is closely related to the field commonly known as chaos theory. We have done this because the word *chaos* has been something of a misnomer, leading to some confusion about the implications of chaotic behavior. The chief source of confusion comes from what has perhaps been an overemphasis on one particular aspect of chaos, sensitive dependence on initial conditions, especially in the popular literature on the subject. This particular characteristic is actually not even uniquely characteristic of strange attractors; unstable points exhibit the same phenomenon. However, the persistence of systems even though chaotic, coupled with the property of sensitive dependence on initial conditions, leads to the intuitive notion that they are inherently unpredictable. Although this is true in a narrow technical sense, it is certainly not the most important aspect of strange attractors (a better name than chaotic attractors).

Consider, for example, a tornado (Vandermeer and Yodzis 1999). That is most likely a chaotic object, an example of a strange attractor. It represents sensitive dependence on initial conditions in the following sense. Consider two particles of dust within the tornado. If they are very close to one another at one point in time, that fact has no bearing on where they will be with respect to one another in the near future. And how close they are now is not at all correlated with how close they will be in the future. So the future location of each dust particle is dependent on exactly where it is now and may change very dramatically with only a very slight change in its position now. That is sensitive dependence on initial conditions. But in fact that is not the interesting thing about a tornado. It is shaped like a funnel—that is, it has a morphology—and the whole thing travels along the ground wreaking devastation wherever it goes—that is, it has a behavior. Furthermore, if you see one coming toward you it is really quite a good idea to get out of the way,

even though it is chaotic and therefore “unpredictable.” Sensitive dependence on initial conditions, the key idea of the unpredictability of chaos, refers only to the behavior of those dust particles inside the tornado. What is truly interesting about a tornado, what we wish to know and even predict, is perfectly knowable and predictable—thankfully. As was so eloquently stated in a recent summary of complexity theory:

First, control of natural phenomena begins to slip out of the grasp of observers, both because sensitivity to initial conditions severely limits the possibilities for prediction and control and because emergent properties of complex systems are unpredictable from a knowledge of parts. . . . Second, these emergent properties can nevertheless be made intelligible in terms of appropriate descriptions of the processes involved, by using high-level concepts that capture their essential aspects. (Solé and Goodwin 2000, 27)

It is probably not of particular interest to determine whether a system is “formally” chaotic or not. An extremely complicated periodic attractor is, for all practical purposes, equivalent to a strange attractor anyway. If the behavior of the system is bounded and there is an instability within the bounded region, for all practical purposes it may be treated as if it were a strange attractor. Granted that there are cases of rather simple periodic attractors that can be analyzed in a traditional fashion. But many of the system behaviors that we can expect of populations embedded in ecosystems are likely to be very complicated, more like those of strange attractors than those of simple points or limit cycles. This does not mean that they cannot be understood any more than it means that a tornado has no shape. The focus should be not on the unpredictability but rather on the morphology of the attractor—where are its boundaries, is it periodic-like, does it have dense and less dense regions, what is its overall shape, and so forth—the “appropriate descriptions” referred to in the above quotation. Much as we find in studying the morphology of organisms, there is no one defining feature of the morphology of strange attractors.

In studying the morphology of attractors, we have presented what seem to be some key principles. Where are the boundaries of a regional attractor? Is it structurally stable (in a practical sense)? What are the natures of nearby bifurcation points? Where are the basins of attraction? These and similar questions are likely to be the ones we are able to answer about ecosystems in general and, we furthermore suggest, are the ones we are more interested in answering anyway, in the pursuit of understanding and/or designing ecosystems.

A bacterial population increases exponentially, at least for a short period of time, on a nutrient agar substrate. A population of rodents is maintained in a region but with dramatic shifts in numbers from year to year, in patterns seemingly like the chaotic patterns reflected in some of the simple models already discussed. These are patterns in time. Acacia trees in the African savannahs tend to occur in thickets where hundreds of individuals are concentrated in relatively small areas, and between these areas of concentration it is uncommon to find even a few individuals (usually the concentrated areas are near areas where water is close to the surface). This, too, is a pattern, but in space rather than in time. In recent years the analysis of this sort of pattern has become increasingly important in population ecology, and extremely sophisticated tools have evolved to aid in this analysis (see, for example, Fortin and Dale 2005). Here we deal with only the most elementary aspects of spatial analysis. The more complicated analyses are indeed quite complicated and beyond the intended scope of this book.

Rather than expand on what are essentially measurement methodologies (the more complicated analyses mentioned above), we include in this chapter an introduction to different ways of thinking about space. First, we explore different ways of describing spatial patterns and the importance of considering spatial scale. Second, we explore the general question of where spatial patterns come from in the first place, focusing on one generalization that may frequently apply, what we refer to as the Turing effect. Finally, we explore an idea that has become the subject of a whole field of ecological research unto itself—that of metapopulations. When populations are highly clumped, with dense patches in a matrix of rarity, one begins thinking at two levels: (1) the level of the analysis of the population dynamics of one of

the patches, which is essentially what has been presented in previous chapters, and (2) the level of the analysis of what happens between patches, considering the collection of patches as the population. This latter focus is the subject of metapopulations.

EXERCISES

- 5.1 Choose 75 random numbers between 0 and 400. Do it again, and pair the numbers with those of the first set. Imagine that these pairs of numbers are coordinates of trees in a plot in a forest (400×400 meters), and graph the points (i.e., make a map of the locations of the trees). Place a grid of 100×100 hectares on the map, count the number of trees in each hectare, and make a graph of the relative frequency of different counts. Now group the hectare classes by threes (i.e., combine the frequency of hectares that have 0, 1, or 2 trees into a single class, the hectares that have 3, 4, or 5 trees in a distinct class, etc.), and recalculate the relative frequencies. What does the comparison of the frequency distribution when grouped by ones versus threes tell you?
 - 5.2 In a forest plot in Michigan measuring 400×400 meters we find 72 American basswood trees with the x and y coordinates as given in the appendix to this chapter. Make a graph of the x coordinates against the y coordinates, and qualitatively compare the pattern you see with the pattern you generated in exercise 5.1. Group the hectare classes by ones (i.e., the frequency of hectares that have 0, the frequency with 1, the frequency with 2, etc.). Group the hectare classes by threes (i.e., the frequency of hectares that have between 0 and 2 trees, the frequency that have between 3 and 5, the frequency that have between 6 and 8, etc.). Compare the relative frequencies for the groups classed by ones and that classed by threes.
 - 5.3 In the same forest plot of 400×400 meters we find 61 American beech trees with the x and y coordinates as given in the appendix to this chapter. Make a graph of the x coordinates against the y coordinates, and qualitatively compare the pattern you see with the pattern you generated in exercise 5.1. Group the hectare classes by ones and then by threes, and recalculate the relative frequencies in both cases.
 - 5.4 Plot the relative frequencies in the data from exercises 5.1–5.3 on the same graph, and compare the patterns.
 - 5.5 In the case of the basswoods in exercise 5.2, there are two particularly large individuals in the population at $x = 178$, $y = 215$ and at $x = 200$, $y = 114$. Replot the data of exercise 5.2, and plot these two individuals with a larger symbol on the plot. Could the positions of these two individuals explain anything about the reason for the spatial pattern?
 - 5.6 In the case of the beeches in exercise 5.3, there is one particularly large individual at $x = 173$, $y = 321$. Replot the beech data of exercise 5.3, and add this individual using a different symbol. Could the position of this individual explain anything about the reason for the spatial pattern?
-

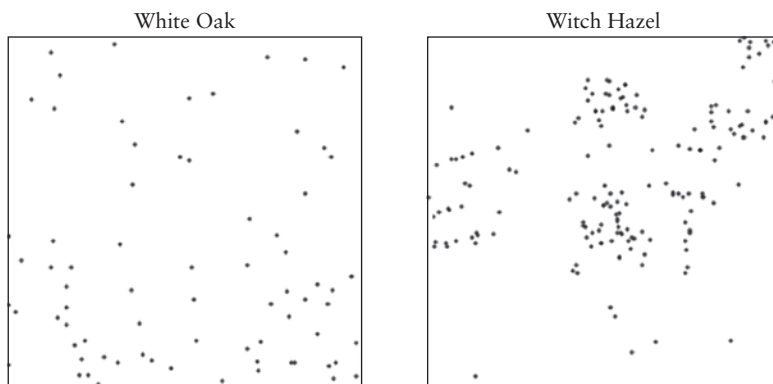


FIGURE 5.1. Distribution of two species of trees on a 1-hectare plot at the E. S. George Reserve near Ann Arbor, Michigan.

We begin with some simple spatial concepts. Consider the illustration in figure 5.1 of the distribution of white oak trees in a 1-hectare plot at the E. S. George reserve north of Ann Arbor, Michigan. There seems to be little pattern to the data; it seems almost like the pattern you might get from a shotgun blast. Also in figure 5.1 is the distribution of witch hazel trees in the same plot, illustrating an obviously different pattern. The witch hazels are clumped up in a couple of different parts of the plot, whereas the white oaks are dispersed around the entire plot. How should such differences be quantified? Are there different spatial scales at which quantification should be focused? What are the dynamic consequences of the difference? We start with the simple problem of how to measure such spatial patterns.

Consider the pattern of birds on a telephone line. At one extreme are the parakeets, who love being together and tend to clump on the line (see figure 5.2A). At the other extreme are the blackbirds, who continually fight with one another and wind up evenly distributed on the line (see figure 5.2B). What sort of index will reflect the obvious difference between these two patterns?

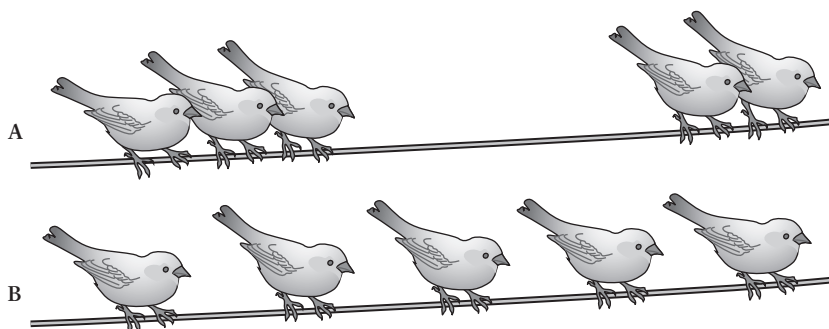


FIGURE 5.2. Birds on a telephone line. (A) Clumped. (B) Evenly dispersed.

The Poisson Distribution

The usual way of approaching this problem is by assuming a sort of null model, at least initially. Let's say that the birds land on the line at random and we measure their positions before they either clump up (as do the parakeets) or disperse themselves (as do the blackbirds). We divide the telephone line into units so small that only a single bird can fit in a unit. There is some probability that a given unit will be occupied (mainly determined by the population density and behavior of the birds in the vicinity). That probability is p . We now ask the question, What is the probability that exactly r units will be occupied on a length of telephone line that contains n spaces (remember, at this point each space can contain only a single bird). For example, suppose that there are only three spaces and $p = 0.4$. What is the probability that exactly two of the three spaces will be occupied? All possibilities are shown in figure 5.3.

In other words, the probability that two of the three spaces on the telephone line will be occupied is $(3)(0.4)^2(1 - 0.4)$. We can generalize this idea by noting that the probability that r of n spaces will be occupied if p is the probability of such space being occupied is

$$Cp^r(1 - p)^{n-r},$$

where C is a constant equal to the number of ways r things can fit into the n locations. In general, the constant C is given as

$$\frac{n!}{(n - r)!r!}.$$

In the preceding example, $r = 2$ and $n = 3$, which gives us

$$\frac{n!}{(n - r)!r!} = \frac{(3)(2)(1)}{[(1)][(2)(1)]} = \frac{6}{2} = 3.$$

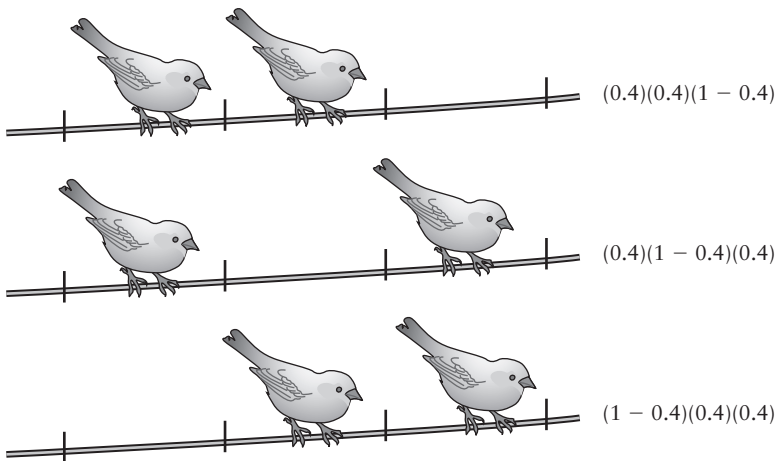


FIGURE 5.3. All possible ways that two birds can fit on three places on a telephone line.

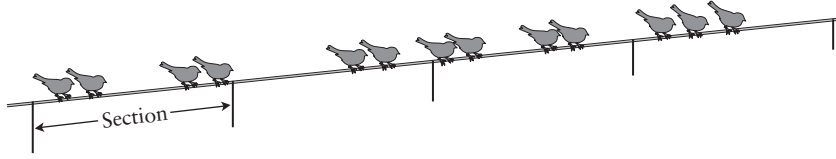


FIGURE 5.4. The telephone line divided into segments (sections), each of which could, theoretically, contain a large number of birds.

Thus, in general, the probability that exactly r locations will be occupied out of a possible n for which the probability of being occupied is p is

$$P_r = \frac{n!}{(n-r)!r!} p^r (1-p)^{n-r}, \quad (1)$$

which is the binomial probability, the r th term in the expansion of $(1-p)^n$.

Now suppose that we have a very long telephone line and we divide it up into relatively large sections (so that some large number of birds could fit into each section). This would look something like figure 5.4.

But we know that each of these segments is made up of a large number of small segments, each of which can contain a single bird. The probability that exactly r of the subsections in a given segment will be occupied by a bird is the same as the probability that the segment will contain r birds, and we already know that that is given by the binomial. However, now we are assuming that there is a very large number of subsections in a given segment of the telephone line. Effectively this means that n is very large compared to r . And this is the trick. Assuming that n is very large with respect to r enables us to derive an exact formula to describe a random distribution (assuming that each of the r subsections is occupied at random for all of the segments is effectively the definition of a random distribution of birds on the telephone line). From equation 1, allowing n to be very large with respect to r , we have

$$P_r = \frac{n!}{(n-r)!r!} p^r (1-p)^{n-r} = \frac{n(n-1)(n-2) \dots (n-r+1)}{r!} p^r (1-p)^{n-r}.$$

From this we can see that $n-r$ (the exponent in the last term) is approximately equal to n (because n is very large relative to r), and $n(n-1)(n-2) \dots (n-r+1)$ is essentially the same as n multiplied by itself r times (because n is so large in comparison to r). Thus we have

$$\frac{n^r p^r (1-p)^n}{r!}.$$

Note that we can define the number of elements (birds) in a section as the number of segments (n) times the probability that each segment is occupied (p), or np . Symbolizing np with λ , we obtain

$$\frac{\lambda^r \left(1 - \frac{\lambda}{n}\right)^n}{r!}.$$

Recall from basic calculus that

$$\lim_{n \rightarrow \infty} (1 - x/n)^n = e^{-x}.$$

Thus, if we allow n to become very large,

$$P_r = \frac{\lambda^r e^{-\lambda}}{r!}, \tag{2}$$

which is the equation for the Poisson series, where λ is the mean number of birds per section (or, more generally, the mean number of elements per quadrat).

EXERCISES

- 5.7 Using the Poisson formula, graph p^r as a function of r for $\lambda = 0.5, 1.0$, and 2.0 . Once you get the formula set up, experiment with other values of λ . What happens to the shape of the function relating p^r to r as λ increases, and why might that happen?
- 5.8 From the data in exercise 5.2, compute the average number of trees per quadrat and apply the Poisson distribution (use the average number of trees per quadrat as the estimate of λ). Graphically compare the Poisson distribution with the data.
- 5.9 From the data in exercise 5.1, compute the average number of trees per quadrat and apply the Poisson distribution (use the average number of trees per quadrat as the estimate of λ). Graphically compare the Poisson distribution with the data.

Applying this formula to the original data on trees at the E. S George Reserve, we redraw the plots with grid lines, as in figure 5.5. Now count the number of quadrats that have 0 trees in them (54 in the case of witch hazel), then the number of quadrats that have 1 tree in them (13 in the case of witch hazel), then the number that have 2 trees, and so forth. The total number of trees in all the quadrats divided by the number of quadrats (in this case 100) gives the value of λ (the mean number of trees per quadrat). With this estimate of the mean, we

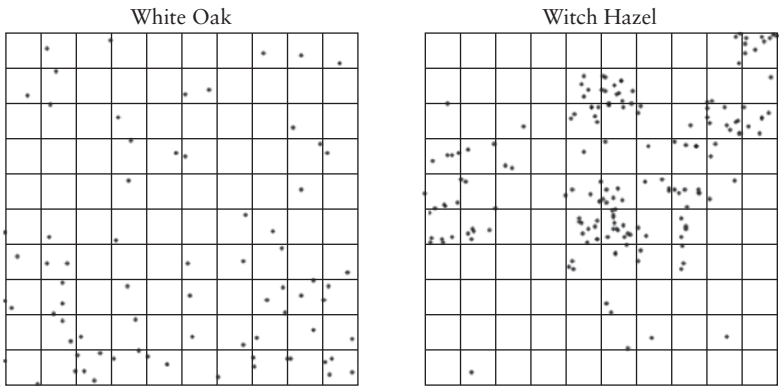


FIGURE 5.5. The same graph as in figure 5.1 but with 10-meter grid lines dividing the plot into 100 quadrats of 10×10 meters.

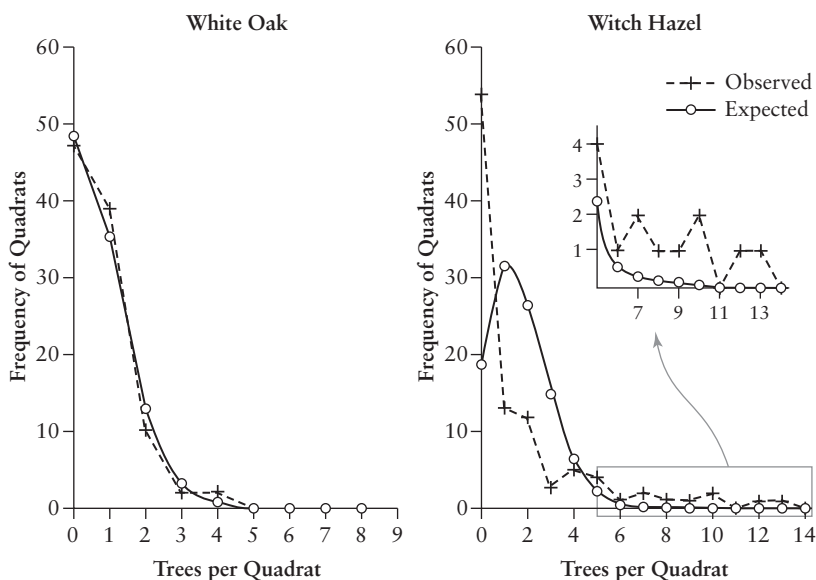


FIGURE 5.6. Comparison of expected (based on the Poisson distribution) with observed density per quadrat of witch hazel and white oak trees in the plot at the E. S. George Reserve. The inset on the graph for witch hazel is a blowup of the right-hand side of the graph, showing the obvious difference between expected and observed.

apply equation 2 and multiply each probability by 100 (the total number of quadrats) to get the expected number of quadrats for each category of tree density. We then compare the expected with the observed, as is done graphically in figure 5.6. As can be seen, the observed values for witch hazel are significantly different from the expected, but the plot for white oak is very much as expected from the random hypothesis (i.e., it more or less follows what would be expected for a random distribution, which is to say a Poisson distribution). Testing the hypothesis of randomness is thus simply a matter of using a chi-square or similar statistic to compare the expected distribution with the observed.

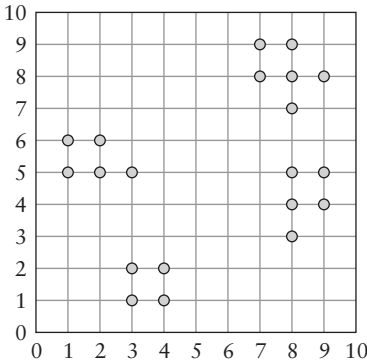
The nature of the witch hazel distribution, obvious from a simple examination of its distribution (e.g., see figure 5.4 or 5.5), is also obvious when comparing the Poisson expectation with the observed. According to the Poisson distribution, we expected 19 quadrats with zero individuals, whereas we observed 54 (see figure 5.6). We expected 31 quadrats with a single individual, but we observed only 13 (see figure 5.6). And lumping all quadrats that contained more than 4 individuals (i.e., really dense clusters of individuals), the random expectation tells us to expect a total of 3 quadrats with these numbers, whereas the actual number of quadrats with more than 4 individuals is 13. Thus we see major deviations of the expected distribution from the observed in the case of witch hazel—the first term of the Poisson distribution (i.e., the term that gives us the probability of observing zero individuals in a quadrat) is way less than the observed (19 as opposed to 54), whereas the final terms (the probability of obtaining more than 4 individuals in a quadrat)

also give us way fewer than observed. This pattern is reflective of what is clear from the original picture (we chose this example because it is so obvious).

The Poisson distribution has the extremely useful property that in a perfectly random distribution, the mean and the variance of the number of individuals per quadrat are identical. Consequently, one can use the ratio of variance to mean as a measure of “clumpiness.” So, taking the example of trees in the E. S. George Reserve, for white oaks the mean number of individuals per quadrat is 0.73. The number of quadrats with 0 trees is 47, the number with 1 tree is 39, the number with 2 trees is 10, the number with 3 trees is 2, and the number with 4 trees is 2. There are no quadrats with 5 or more individuals. So the variance will be $47(0.73 - 0)^2 + 39(0.73 - 1)^2 + 10(0.73 - 2)^2 + 2(0.73 - 3)^2 + 2(0.73 - 4)^2 = 75.71$, divided by the number of quadrats less 1 (99), which gives 0.765. The variance-to-mean ratio is then $0.765/0.73 = 1.05$, which is very close to the theoretical expectation of 1.00 for a random distribution. Making the same calculation for witch hazels, we find a mean of 1.68 and a variance of 8.76, making the variance-to-mean ratio equal to 5.21. Thus we conclude that the spatial distribution of white oaks is consistent with a random expectation, whereas the witch hazels are highly clumped. Furthermore, the variance-to-mean ratio may be thought of as a quantitative measure of the “degree” of clumpiness. A ratio greater than 1 indicates a clumped distribution, a ratio of less than 1 indicates an evenly dispersed distribution, and a ratio equal to 1 indicates a random distribution.

EXERCISES

- 5.10** Consider the following two artificial populations distributed in space, with coordinates 1,1; 1,3; 2,2; 2,1; and 9,9 for the first population and 1,5; 5,3; 6,6; 8,9; and 9,1 for the second population. Make maps for both of the populations separately. What is the difference between the two? Now use the Pythagorean theorem to compute all the distances between pairs of points for each of the species (in each case, there are 10 such distances). What is the mean distance for each of the populations? What is the variance? Order each set of distances independently, and plot together on a rank axis (i.e., smallest distance first, next smallest, next smallest, . . . greatest distance).
- 5.11** Consider the following artificial distribution of an organism in a space of 10×10 meters:



Considering interindividual distances of 1.0 or less, by eye count, how many distances are there, and what is their average? Considering interindividual distances of 1.5 or less, by eye count how many distances are there (remember that the square root of 2 is about 1.4) , and what is their average? Repeat for distances of 3.6. Estimate approximately (by eye) the average distances between the two pairs of clumps. Finally, estimate approximately the average distances between the two large clumps. Make a graph of the average distances as a function of the critical interindividual distance. What do you think this may have to do with the clustering pattern?

Point Pattern Analysis and the Question of Scale

Dividing an area into quadrats and counting the number of individuals in those quadrats (as we did above) is especially useful when the data gathered are in the form of number of individuals per quadrat. Calculating the mean number of quadrats with 0, 1, 2, . . . individuals and comparing it to a Poisson expectation is a straightforward task whose results are easily interpreted. However, when data are available in a georeferenced format—which is to say that each individual in a plot has an x and a y coordinate associated with it—other methods are available, broadly categorized as “point pattern analysis.” In ecology the basic idea was noted long ago in a critical article (Clark and Evans 1954) in which the idea of nearest-neighbor analysis was introduced. It is a simple and intuitively attractive idea, as illustrated in figure 5.7. The white oaks, partly because the local population density is less dense, have larger distances between nearest neighbors, whereas the witch hazels, partly because the population density is larger, has very small distances between nearest

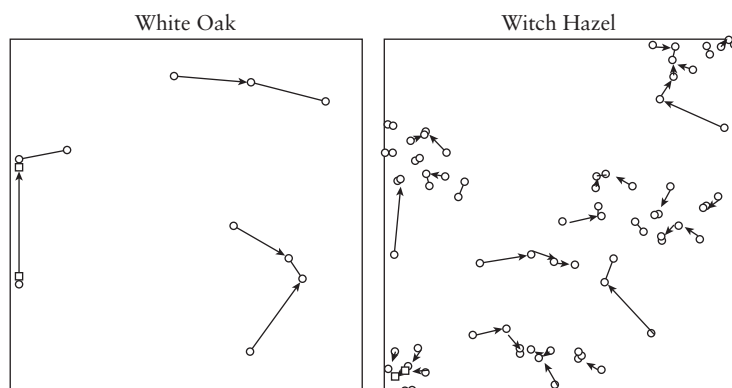


FIGURE 5.7. White oak and witch hazel plots, the upper right quarters of the plots shown in figures 5.1 and 5.5. Lines connect the nearest neighbors, where lines without arrowheads connect both trees (i.e., the direction of connection goes both ways) and lines with arrowheads connect a single individual to some other individual. Thus each tree (each point in the quadrat) is “nearest” to some other tree. It is evident that, on average, the distance to nearest trees in the case of white oak is large, whereas the distance to the points in the case of witch hazel is much smaller.

neighbors. Clearly the average nearest-neighbor distance reflects both population density and the degree of clustering of the population. Clark and Evans (1954) basically took the population density effect out of the equation by comparing the actual nearest-neighbor distances to what would be expected from a random spatial distribution of individuals of that same density.

The concept of nearest-neighbor distances can be expanded to look at the distance between a single individual and all of its neighbors in a given area as another way of using point- rather than quadrat-level data to analyze spatial patterns. For any individual and neighborhood area, we can calculate the actual mean distance to all its neighbors and the deviation between this observed value and that expected if individuals were distributed at random in that neighborhood. The mean value of the summation of these deviations is known as Ripley's K and is extremely popular as a way to reduce a complex spatial pattern to a simple quantitative representation. Typically a Ripley's K analysis is repeated for a wide range of neighborhood areas or scales and presented as the ratio of average nearest-neighbor distances to what would be expected with that number of neighbors if they were randomly allocated within that scale. A stylized version of a Ripley's K analysis is presented in figure 5.8, where it is clear that no generalization can be made about the nature of a spatial distribution without specifying the scale at which the observations are being made.

What is evident from the above is that the scale at which observations are made can strongly affect the nature of the observed pattern. Another way to

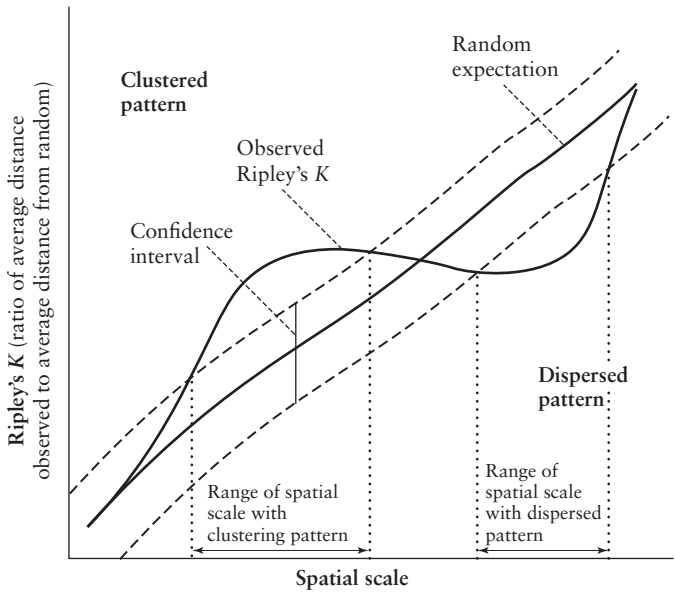


FIGURE 5.8. Stylized version of Ripley's K analysis, showing Ripley's K as a function of spatial scale. Whenever K is greater than the 95% confidence limits, the spatial pattern can be said to be clustered. Whenever K is less than the 95% confidence interval, the spatial pattern can be said to be dispersed.

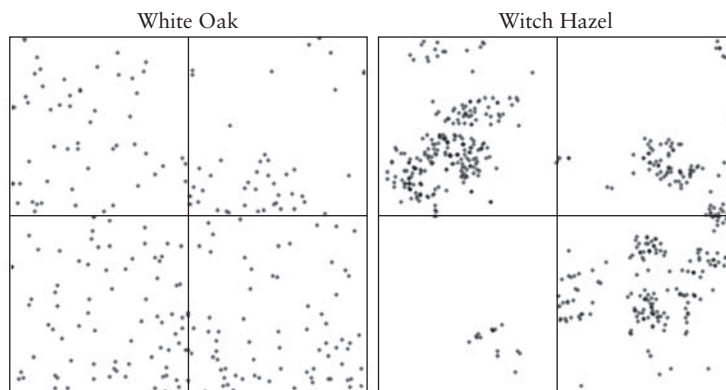


FIGURE 5.9. Spatial distribution of two tree species at the scale of 4 hectares.

make this very same point is to simply examine different scales. For example, for the forest from which the above running example has been taken, if we double the sides of the sampling region and look at four contiguous hectares rather than the single one or the quarter of one that we have been working with, we see the emergence of an even more complicated pattern (figure 5.9). What seemed to be clusters at the scale of a hectare seem to be clusters within bigger clusters at this larger scale for witch hazels, and the apparent random distribution of white oaks seems to have developed a nonrandom component, with one hectare (the upper right) distinctly nonrandom. When we move our lens out further (figure 5.10), the witch hazels seem to have different-sized clusters of different “textures,” and the white oaks still seem to suggest a non-random pattern (although it is actually not significantly different from random). In the end, no matter whether we focus ever more closely or ever more distantly, the scale of observation makes a difference in the conclusions we are likely to reach.

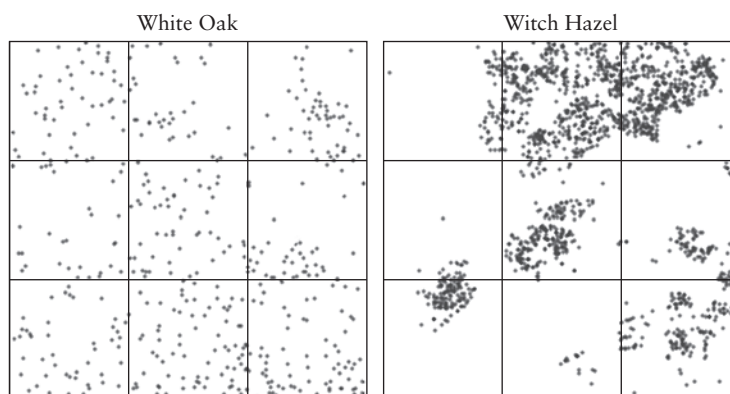


FIGURE 5.10. Spatial distribution of two tree species at the scale of 9 hectares.

Mechanisms of Spatial Pattern Formation: Principles of Reaction/Diffusion

In this chapter we have so far emphasized the process of making observations about and statistically analyzing the patterns one might see when observing populations in nature. But the question of where those patterns might come from has not yet been discussed. Indeed its answer is one that, on the one hand, would seem somehow obvious but yet, on the other hand, is especially enigmatic. Reminiscent of the early debates (see chapter 2) about density independence versus density dependence, the issue of spatial pattern can be cast in a similar way. The obvious reason for clustered spatial patterns in organisms is simply that they reflect the clustering of their habitats. For example, a beetle that specializes on the acacia trees in the African savannah that we mentioned at the beginning of this chapter would likely be clustered if we looked at a map of the points where it occurs on a large map. But those clusters would simply correspond to the clusters of the acacia trees (because the beetle could not live anywhere else). So one might ask what causes the acacia trees to be clustered, and in this case, we know that those clusters tend to form in patches of land that happen to have a slightly higher level of soil moisture. Thus in both cases (beetle and acacia) the clustered spatial pattern is a product of some underlying habitat variable that itself is clustered.

More interesting, perhaps, are the many circumstances in which a pattern occurs even though the background habitat is uniform. This sort of pattern formation is sometimes referred to as autonomous pattern formation because there is something within the population dynamics that causes that pattern. In figure 5.10, for example, each of the small clusters of witch hazel trees stems from the seed dispersal of one or two “seed source” individuals located in that cluster. So the observed clusters are a consequence of a biological process inherent to the population in question, namely, seed dispersal. Whether an observed pattern comes from some underlying pattern independent of the dynamics of the population or whether it comes directly from those dynamics worked out in space is reminiscent of the older debate about whether a population is regulated through forces independent of its density or is dependent on its density.

The pattern of individuals that represent the seed shadow of one or two isolated individuals is an obvious way to generate a clustered pattern, but it also points to a second generalization: if dispersal is the only process operative, clustered patterns will be observed only over a short time period and are, inherently, transient. Are there circumstances in which the inherent population dynamics of a population will result in a spatial pattern, even if the background habitat is completely uniform, and that pattern will be persistent over time?

One very general approach to this problem is informed by the work of the British mathematician Alan Turing (1952). Although Turing thought of his work as applying to chemicals in space, its application to populations in broad general terms is quite obvious, as we describe presently. There are basically two processes involved, activation (because the ideas were developed in

the context of chemical reactions, the activation part is frequently referred to as reaction) and diffusion, and thus the equations describing the process generally are referred to as reaction/diffusion equations. We first look at the simple process of diffusion.

EXERCISES

- 5.12** A deranged millionaire is transporting a population of poisonous snakes in a truck with the intent of transporting them to another location on a flight out of Detroit. Only 100 miles from the airport, the truck carrying the poisonous snakes tips over. It is estimated that the probability of snakes' dispersal over a 10-mile distance is 30% per day. If the population was 10,000 originally, what will the dispersal process look like? Use a discrete approximation of the process, and suppose that you are sampling along the highway, so you can assume a one-dimensional space. Because the snakes move in both directions, the equation for the number of snakes at spatial position x is

$$N(x) = N(x) + 0.3[N(x-1) + N(x+1) - 2N(x)],$$

where $x = 1$ represents 10 miles. What will the spatial distribution of snakes look like over time? (Begin with 1,000 in a column of a spreadsheet, type the above equation in the cell below, fill in the cells to the right and left, and then fill in the rows down.) Graph the spatial pattern 12 days into the future.

- 5.13** A situation similar to that in the previous exercise exists, but this time an active bacterial bioweapon is being transported with the intent to infest the entire University of Michigan community. Not only do the bacteria disperse at a rate of 30% per day, they also increase their population at an exponential rate. Repeat exercise 5.12 but with the exponential growth process operative. Use an exponential growth rate of 1.1 and compare the dispersal pattern in space to one with a growth rate of 1.3.
-

We begin with a one-dimensional space (i.e., think of a shoreline species of plant along the edge of a lake or of a roadside weed along a highway). We begin with 100 individuals at one point on the shoreline and proceed with diffusion away from that point. Supposing that the diffusion rate in a given direction is 10%, we get the pattern shown in figure 5.11.

We began with a population of 100 individuals at spatial position 6, and after one unit of time, 20% of those individuals diffused to the neighboring spatial positions of 5 and 7, half to 5 and half to 7. In the next time interval, again 20% of the individuals at spatial positions 5, 6, and 7 each moved to the neighboring positions (from position 5 to 4 and 6, from position 7 to 6 and 8). It should be clear that the population eventually disperses over the entire landscape.

It is possible to generalize this process with a bit of somewhat complicated, although very intuitive, mathematics. Consider the situation of spatial position 7 at time 3 in figure 5.11. That position at that time contains 16 individuals, the position to its left has 66 individuals, and the one to its right has 1 individual. As we move to time 4, the population density increases by 3.5 to 19.5 individuals. That is, the derivative with respect to time (dN/dt) had to

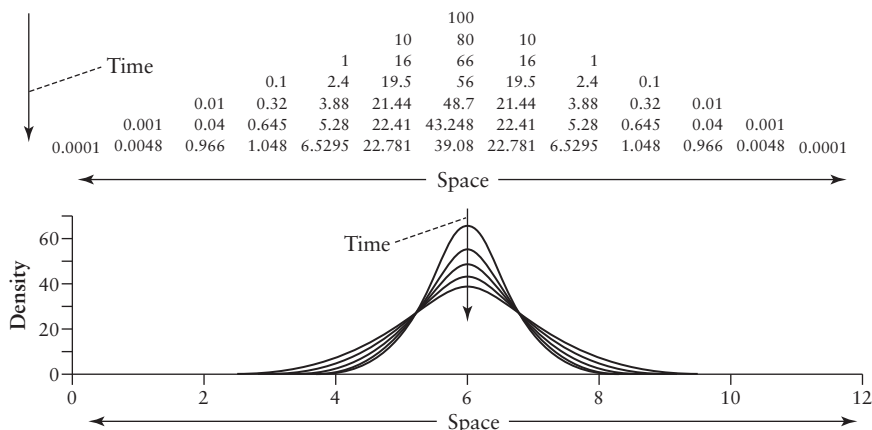


FIGURE 5.11. Artificial example of the diffusion of individual plants in a one-dimensional space. Beginning with 100 individuals at spatial point 6, 10% of the individuals migrate (diffuse) to nearby points in each direction at each time step. Note that here the process is presented as operating in discrete space and discrete time.

be positive, indeed had to be about 3.5 (the difference between 16 and 19.5). Where did that number come from? The rule we used to create figure 5.11 is to allow individuals to migrate (diffuse) to neighboring spatial positions at a rate of 10% in a given direction. So 20% of the 16 individuals, or 3.2 individuals, would have to leave position 7. But 10% of the individuals from each of the two neighboring positions would also have to go to position 7, which is to say that 6.6 individuals from the left (10% of 66) and 0.1 individuals from the right, or a total of 6.7 individuals, would have to arrive at position 7. Subtracting 3.2 from 6.7 gives us the required 3.5 individuals. Note that we calculated the change in population density over time with reference only to changes that were happening in space.

We can formalize the above numerical examples. At any particular point, x , there will be $N_x(t)$ individuals at time t . The number of individuals at time $t + 1$ will be that number minus those that diffused away, plus those that diffused in, or

$$N_x(t + 1) = N_x(t) + DN_{x-1}(t) + DN_{x+1}(t) - 2DN_x(t), \quad (3)$$

where D is the rate of diffusion or the proportion of individuals moving in one direction from a point. At this point we wish to use equation 3 heuristically, perhaps violating formal mathematical procedures but providing an intuitive grasp of what the mathematics of diffusion actually means. Let us change equation 3 to read

$$N_x(t + \Delta t) - N_x(t) = D\{(N_{x-\Delta x} - N_x) - (N_x - N_{x+\Delta x})\},$$

where $\Delta t = \Delta x = 1$ in equation 3 and we have left off the functional notation of time (t) in all the N s on the right-hand side of the equation (so it will not look so complicated). If we now allow Δt and Δx to approach zero, we see that

$$N_x(t + \Delta t) - N_x(t) = \frac{\partial N_x}{\partial t} \Delta t$$

and

$$N_{x-\Delta x} - N_x = -\frac{\partial N_x}{\partial x} \Delta x,$$

so we can write

$$\frac{\partial N_x}{\partial t} = D \left(\frac{\partial N_x}{\partial x} \Big|_{x-\Delta x} - \frac{\partial N_x}{\partial x} \Big|_{x+\Delta x} \right).$$

It is critical to note that the term in the brackets is the difference between the two partial derivatives evaluated just to the left and just to the right of the critical point x . In general, the difference between two derivatives is the second partial derivative, so the general equation becomes

$$\frac{\partial N_x}{\partial t} = D \frac{\partial^2 N_x}{\partial x^2},$$

which is the classic equation for diffusion.

To this point we have looked only at diffusion. If we now add exponential population growth to the system, equation 3 becomes

$$N_x(t + 1) = N_x(t) + \lambda N_x(t) + DN_{x-1}(t) + DN_{x+1}(t) - 2DN_x(t). \quad (4)$$

Following the above logic of allowing Δt and Δx to approach zero, we obtain

$$\frac{\partial N_x}{\partial t} = \lambda N_x + \frac{\partial^2 N_x}{\partial x^2},$$

which is the equation for diffusion with population growth. To make the picture more general, we note that the dynamics of the population need not be precisely exponential, and to be as general as possible we can simply describe the growth of the population as $f(N_x)$, which gives us

$$\frac{\partial N_x}{\partial t} = f(N_x) + D \frac{\partial^2 N_x}{\partial x^2}, \quad (5)$$

which is the classic form of the reaction/diffusion equation (the “reaction” is $f[N_x]$, which can be any rule of local population growth).

As a specific example, suppose we use the logistic map instead of exponential growth. Then, using equation 4 as a starting point, we can write

$$N_x(t + 1) = N_x(t) + \lambda N_x(t)[1 - N_x(t)] + DN_{x-1}(t) + DN_{x+1}(t) - 2DN_x(t). \quad (6)$$

(For a more analytical approach, the interested student can view the original research of Fisher [1937], who used the partial differential equation form of equation 6 to describe the spread of genes in spatially distributed populations.) If we use this equation to generate the population over time and space (beginning with 90 individuals and allowing the logistic term to have a carrying capacity of 100), we obtain the graph shown in figure 5.12, which is parallel to the one shown in figure 5.11.

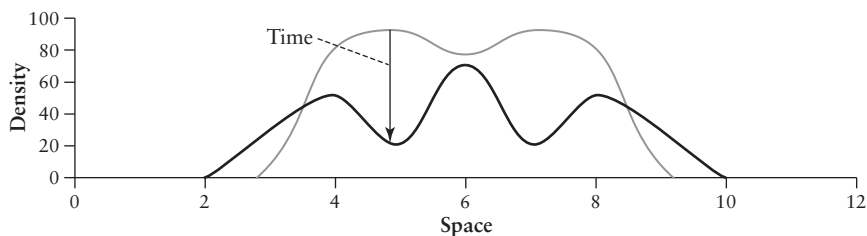


FIGURE 5.12. Spatial distribution of individual plants after 6 (light curve) and 7 (boldface curve) iterations in time, using equation 6 (but with $rN_x((100 - N_x)/100)$ rather than scaling the carrying capacity to 1.0 and with $\lambda = 3.9$, i.e., a chaotic population). Note that the boldface curve has “clusters” of individuals at spatial points 4, 6, and 8.

We thus see the general conclusion that if the reaction term, that is, the population growth term, of the reaction/diffusion equation is nonlinear, a spatial pattern can emerge in an autonomous fashion, which is to say, as a consequence of the dynamics of the population itself rather than that of any sort of underlying heterogeneity of the environment (the environment across the spatial gradient from location zero to location 12 in figure 5.12 is completely homogeneous). The important point is that, in the context of simple reaction/diffusion systems, if the reaction term is nonlinear, it is possible to generate a spontaneous spatial pattern above and beyond the smooth localized pattern normally thought of in diffusion systems (the drop of ink in a bucket of water). In a sense this result is not at all surprising because we simply take the consequences of nonlinearity in the context of discrete time and extend it over space. In the following section we further explore the dynamics of diffusion in two dimensions.

Mechanisms of Spatial Pattern Formation: Biological Causes

We now return to the details of Turing’s famous insight. First, we take the basic reaction/diffusion equation (equation 5) and cast it in two-dimensional space. The space will now be represented by two variables, x and y , rather than only one, and the equation is expanded to

$$\frac{\partial N}{\partial t} = f(N) + D \left(\frac{\partial^2 N}{\partial x^2} + \frac{\partial^2 N}{\partial y^2} \right).$$

This is the standard form of the reaction/diffusion system in two-dimensional space. Turing expanded on the idea that f could be nonlinear (as discussed above) and asked what would happen if there were two distinct reaction terms that interacted with one another in the reaction part of the equation. This means that, much as we can expand the argument from a one-dimensional space to a two-dimensional space, we can expand the reaction variables from one dimension to two dimensions. In the ecological context this might correspond to adding a predator or a parasite to the basic system, among other possible examples.



FIGURE 5.13. Examples of Turing patterns emerging from equations 7a and 7b. The black might be a prey, while the white is its parasite, and the patterns exist in two-dimensional space for three different parameter sets.

The basic equation is duplicated for the second reaction term, obtaining

$$\frac{\partial N_1}{\partial t} = f_1(N_1, N_2) + D_1 \left(\frac{\partial^2 N_1}{\partial x^2} + \frac{\partial^2 N_1}{\partial y^2} \right) \quad (7a)$$

and

$$\frac{\partial N_2}{\partial t} = f_2(N_1, N_2) + D_2 \left(\frac{\partial^2 N_2}{\partial x^2} + \frac{\partial^2 N_2}{\partial y^2} \right). \quad (7b)$$

Turing's insight was to ask what would happen if D_1 and D_2 were distinct from one another and how that might relate to the two species in the reaction part of the system.

Suppose that N_1 is a species that is “autocatalytic,” which is to say that it increases on its own (a growing population, for example). Suppose that N_2 is a parasite or a predator, which is to say that it is some entity that reacts positively to the value of N_1 but then acts negatively on that same variable. Turing referred to N_1 as the “activator” and N_2 as the “repressor.” The basic result is that under the circumstances of an activator and a repressor, if the diffusion rate of the repressor (D_2) is greater than the diffusion rate of the activator (D_1), strong and evident spatial patterns may emerge, examples of which are presented in figure 5.13.

The qualitative nature of Turing's insight is the important part. Its basic operation is illustrated in figure 5.14.

Metapopulations

The construction of classic population models (e.g., those in chapters 1 and 2) tacitly assumes a relatively uniform distribution of individuals in the population such that all individuals in the population interact with and, in the case of a sexual population, can mate with all other individuals in the population. It is the notion of a panmictic population. Yet it is clear to the most casual observer that most populations in nature are patchy in distribution. Habitats tend to occur in a patchwork, and consequently the occupants of those habi-

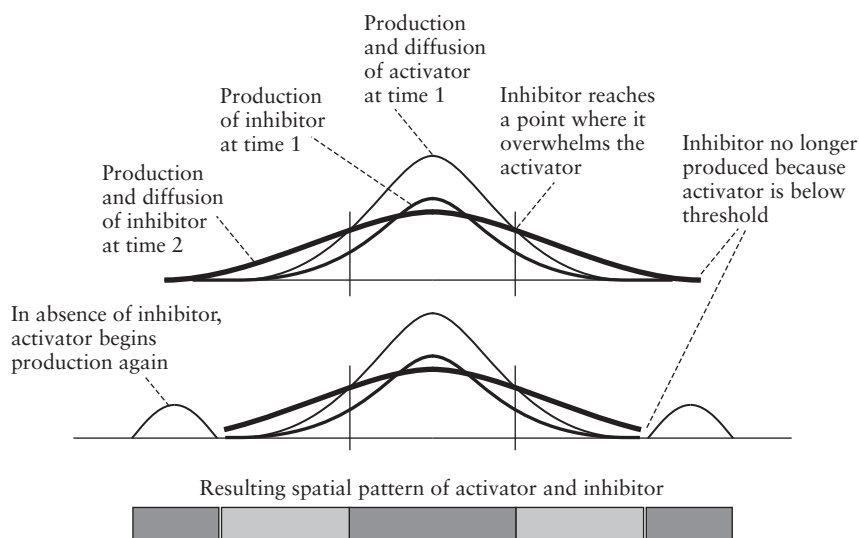


FIGURE 5.14. Diagrammatic representation of the Turing effect.

tats likewise tend to be patchy, whether as a consequence of underlying habitat variables or of the population processes themselves, as in the first part of this chapter. This is obvious in the case of islands, but it is also worth noting that most nonisland populations have something of an islandlike structure when looked at spatially. Interest in the dynamics of such spatially distributed populations has expanded considerably over the past two decades, and the results appear to have a variety of practical applications.

Harrison (1991) provided a useful classification of different forms of spatially distributed populations, a simplified version of which is in figure 5.15. First is the mainland–island (also called source–sink) type, where the sole source of colonists is a large “mainland” or source area in which the population is persistent and colonists from the mainland disperse to “islands” or sink areas in each one of which the population would eventually go extinct if cut off from the mainland source of colonists (figure 5.15A). The next two cases in figure 5.15 include the panmictic population, with a clumped spatial distribution (figure 5.15B), and the metapopulation (figure 5.15C), a term first coined by Levins (1969). The latter general concept has spawned an enormous literature in the past two decades. It has become a springboard for a new way of looking at population dynamics, because it breaks completely with the tradition of looking at birth and death rates of individuals within populations and begins instead with colonization and extinction of entire populations. Much like mainland–island systems, a metapopulation has subpopulations on isolated islands. But rather than experiencing extinction and recolonization only from a mainland or source area, the already occupied islands serve as a generalized source of new migrants. Thus although we expect extinctions to occur on the islands, they are inevitably followed by recolonization from one of the islands that contains an extant population.

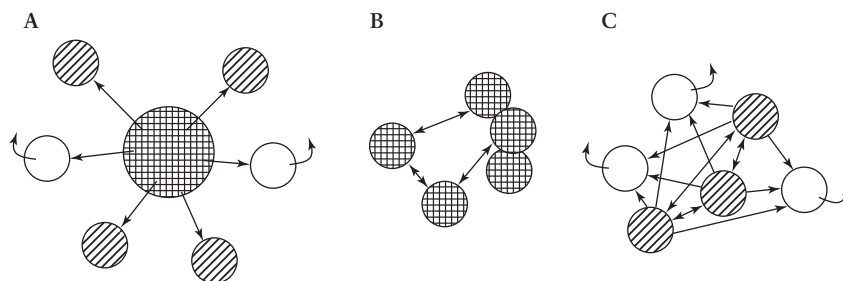


FIGURE 5.15. Simplified classification of spatially structured dynamics (modified after Harrison 1991 and Harrison and Taylor 1997). Circles with diagonal lines represent populations that are destined for extinction (e.g., because their habitat does not offer a full complement of what the population needs to survive) and thus will go extinct eventually. Cross-hatched circles represent populations that are permanent in that they are not ever expected to disappear. Empty circles represent habitats from which populations have disappeared (gone extinct) and are not yet recolonized. Straight arrows indicate migration directions; curved arrows represent previous extinction events. (A) A source–sink population. (B) A panmictic population with a clumped spatial distribution. (C) A metapopulation.

EXERCISES

- 5.14** Remoras are small fish that hang around on the surface of sharks and feed off the remains of whatever the sharks eat. Frequently a given shark can have a large number of these fish on its surface, but sometimes individuals have none. The only way a remora can get from one shark to another is when the sharks come in contact with one another (or get close to one another). If 25% of the individuals in a particular shark population have remoras, what is the probability that a shark without remoras will acquire them upon encountering another shark? (Think of the contact process—what is in contact with what and what needs to happen for an unoccupied shark to become occupied.)
- 5.15** Repeat exercise 5.14 but with the additional knowledge that every time one shark encounters another, there is a 50% chance that a remora (if present) will change sharks.
- 5.16** Let the fraction of sharks with remoras be equal to p and the probability of a remora transferring between two sharks in contact be m . If the rate of change of the fraction of sharks occupied is taken to be equal to the probability of a transfer upon contact, write the general equation for that rate of change.

The theory of island biogeography (MacArthur and Wilson 1967) was perhaps the starting point for thinking about populations dynamically in the context of spatial structure, although that theory dealt with assemblages of species, not with populations. Levins (1969) effectively extended the ideas of island biogeography theory. He first conceived of a population distributed in patches. Key to the idea was that no single patch could form a successful population in perpetuity because all patches had some finite probability of extinction. The cause of extinction is not specified but is assumed to be some stochastic event such as

a major disturbance or arrival of a predator. Thus the only thing that keeps the overall population (the collection of subpopulations over all the patches) alive is migration from patch to patch to allow recolonization of patches in which the population has gone extinct. Rather than constructing the model from the typical state variable of population density, Levins conceived of the overall population as a collection of potential patches, some of which were actually occupied by a subpopulation and others of which were not yet occupied (or reoccupied). Then he speculated that, much as in island biogeography theory, the dynamics of the overall population could be described as the dynamics of the local subpopulations going extinct and being recolonized by individuals migrating from other subpopulations. That is, rather than the birth and death of individuals, we have the colonization and extinction of patches.

Mathematically, the basic Levins model of metapopulations begins by analyzing the state variable p , where p is the proportion of patches currently occupied by a subpopulation. Note that the use of p rather than N (population size or biomass) is a fundamental departure from classic population ecology. If p represents the proportion of habitats occupied, the dynamics of the process can be described as a balance of the probabilities of migration and of local extinction. Initially, let's assume that the extinction rate is zero and p is very small (close to but not exactly equal to zero). The rate of increase of p is then likely to be equivalent to an exponential process because the larger p becomes, the more likely it is that more migrants will be available to occupy empty patches. Thus we can write

$$\frac{dp}{dt} = mp,$$

where m is the migration coefficient. This equation represents an exponential process with p as the dynamic variable and predicts that p should increase exponentially. Because p is a proportion, its maximum value is 1.0, which is analogous to the carrying capacity in the earlier derivation of the logistic equation. As the population approaches its carrying capacity (in this case 1.0), its rate of increase decreases. Indeed we can expect that the migration rate will really be inversely proportional to the proportion of habitats not yet occupied, $1 - p$. Thus the above exponential form really should read

$$\frac{dp}{dt} = mp(1 - p),$$

which is, mathematically, identical to the logistic equation. However, in developing this model we assumed that the extinction rate is zero. If we now add to the model the process of extinction, as if it were an independent death rate in the logistic model, we can write

$$\frac{dp}{dt} = mp(1 - p) - ep,$$

where m is the migration rate and e is the extinction rate. Thus the term for the migration rate is formulated mathematically as a logistic equation (with

carrying capacity equal to 1.0), and extinction is added as a simple linear function.

This simple model generates some important insights. Setting the derivative equal to zero and solving for p , we find the equilibrium state as

$$p^* = 1 - \frac{e}{m},$$

which immediately suggests a couple of interesting ecological conclusions. First, there will be a positive value of p^* (and thus a viable population) whenever $m > e$. This is not really surprising because it simply says that when the migration rate is greater than the extinction rate, the overall population, the metapopulation, will not disappear. The second conclusion is that as long as the extinction rate is greater than zero, the equilibrium state of the population will always include some empty habitats. As a practical matter for conservation, for example, this suggests that a strategy for preserving a species that requires the species to be present in all available habitats may not be realistic. If the population normally exists as a metapopulation, it is inevitable that some habitat patches will always be devoid of individuals of the species, and in fact, preserving empty, suitable patches is an integral part of any successful conservation strategy.

Whether or not a population is truly a metapopulation depends on a variety of criteria (see below), two of the most important of which are that (1) some potential sites are actually unoccupied (see previous paragraph) and that (2) unoccupied sites are potentially occupiable. To examine the second crucial criterion, a variety of experiments have added propagules to sites within a region that are unoccupied but seem suitable. For example, adding seeds to patches of meadows in a forest matrix, workers in Sweden discovered that indeed many of the species could persist in currently unoccupied sites (Eriksson and Kiviniemi 1999). In another study, Tilman (1997) added seeds of up to 54 species to plots in each of 30 prairie openings in native oak savannah. All species occur locally but not in the experimental plots. A significant increase in species richness was found with the number of species added (see figure 5.16). Thus again many of the species seem to be limited in distribution by colonization or propagule availability rather than by site suitability.

Assumptions of Metapopulation Models

Six basic assumptions underlie the idea of metapopulations:

1. Suitable habitat can be conceived of as occurring in discrete patches. This assumption seems to be fairly obvious in many cases, for example, in meadows surrounded by forest or forests in an agricultural matrix.
2. The population is at least temporarily a reproductive population. For example, in a series of dry meadows in a matrix of forest in Finland, Hanski and Simberloff (1997) found that 60%–80% of the butterflies spend their whole lives in their natal patch, providing strong evidence that a patch does not contain just an ephemeral aggregation.

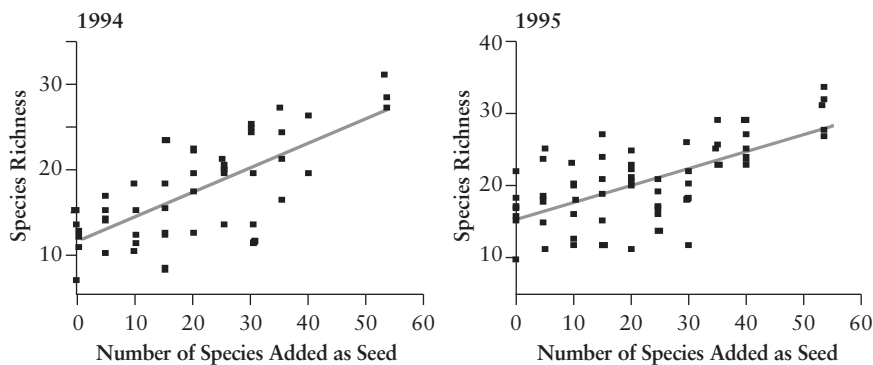


FIGURE 5.16. Results of Tilman's (1997) seed addition experiments (see text).

3. Subpopulations have a substantial risk of extinction. In the Finland butterfly example, the largest population had only about 500 individuals (out of 377 populations), and populations with several hundred individuals have been recorded as going extinct.
4. The subpopulations are not so isolated that recolonization of empty patches is impossible. Using the Finland butterfly example again, the mean nearest-patch distance was 240 meters and the median distance moved by migrating butterflies was 330 meters, more than sufficient to allow colonization to balance extinction. On the other hand, it may be that many fragmented populations are in a state such that colonization rates will not be high enough to sustain true metapopulation dynamics.
5. The dynamics are not synchronized across subpopulations. This is a major concern in applying metapopulation theory to conservation. If extinctions are correlated among patches, a uniform dispersal rate may not be able to compensate for the peaks in extinctions.
6. All patches are alike (i.e., equally suitable) both in terms of probability of extinction and colonization and as a source of colonists. Harrison and Taylor (1997) argue that variation among patches is really the case and that it is likely that most metapopulations are closer to the source-sink model (a source, by definition, has a zero probability of going extinct and also serves as the sole source of propagules) than the classic metapopulation.

Although some of these six basic assumptions have been relaxed in theoretical work of recent years, other aspects may turn out to be as important as research advances. A variety of subjects of obvious importance to the idea of metapopulations have yet to be fully analyzed. For example, variation in the quality of the matrix in which the patches are located will likely be very important but has received little attention (Perfecto and Vandermeer 2010). Although it is rarely acknowledged explicitly, an additional assumption of metapopulation theory is that the matrix is uniformly inhospitable, an assumption that may not be true. The spatial arrangement of patches is

likewise an important issue, as is the nature of the boundaries of the patches. Below we examine two of the complicating issues of metapopulation theory, the rescue effect and the idea of propagule rain.

EXERCISES

- 5.17** Graphically show the general relationship between p^* and e for various values of m .
- 5.18** Some populations may behave as metapopulations in the sense that there is a discontinuous distribution of individuals in space and a constant extinction rate of those patches, but the migration rate is independent of the proportion of habitats occupied (sometimes referred to as propagule rain). Presume that there is zero migration from patch to patch but a constant rain of individuals from outside the system. What would the equation for this situation be? What would the equilibrium value be? Graph the equilibrium as a function of the extinction rate (for propagule rain = 0.25), and compare it to the classic equation with $m = 1$).
-

The Rescue Effect and Propagule Rain

An important assumption in the classical Levins metapopulation model is that the extinction rate is linear with respect to the proportion of sites occupied (i.e., total extinction = ep), which means that a constant proportion of occupied sites is expected to go extinct at every time step. Hanski (1982) took exception to this assumption, reasoning that if p was high it would be likely that a subpopulation that was about to go extinct might be effectively “rescued” from its fate by migration from nearby subpopulations. So the extinction rate itself would be a function of the proportion of sites occupied. Thus Hanski proposed the model

$$\frac{dp}{dt} = mp(1 - p) - ep(1 - p),$$

which can be rewritten as

$$\frac{dp}{dt} = (m - e)p(1 - p),$$

which is the same as the simple logistic equation with the intrinsic rate of natural increase $m - e$ and the carrying capacity $K = 1.0$. Obviously, with a negative intrinsic rate of natural increase (which would be the case if the extinction rate was greater than the immigration rate), the overall metapopulation would tend toward extinction (i.e., p would tend toward zero), whereas if the immigration rate were greater than the extinction rate, the overall metapopulation would tend to fully occupy all the sites (i.e., p would tend toward 1.0). Thus Hanski’s model effectively predicts that there should be a bimodal distribution of populations when metapopulation dynamics are operative. This is referred to as the core–satellite hypothesis. The key here is that a given species, or at least a metapopulation of a species, shifts between these two types, not that there are some species that are core and others satellite.

Gotelli (1991) extended the basic metapopulation model further by suggesting that the quadratic form of the immigration part of the original Levins model was too restrictive in that it presumed that propagules are more likely to arrive when more of the subpopulations are occupied. Although this may be reasonable in some situations, in others (e.g., the case of plants with many highly dispersible seeds) it does not seem to fit with reality. In many situations there is likely to be a “rain” of propagules that is relatively independent of the number of habitats occupied; this characterizes a true island–mainland situation, for example. Gotelli thus suggested modifying the immigration part of Levins’s equation, obtaining the following model:

$$\frac{dp}{dt} = m(1 - p) - ep(1 - p),$$

where obviously the rescue effect has also been included. Here the equilibrium value is

$$p^* = \frac{m}{e}.$$

Both the rescue effect and the propagule rain are illustrated graphically in figure 5.17.

The only other possibility left is seen when there is no dependence on regional occurrence at all, namely

$$\frac{dp}{dt} = m(1 - p) - ep,$$

which combines Gotelli’s immigration and Levins’s extinction.

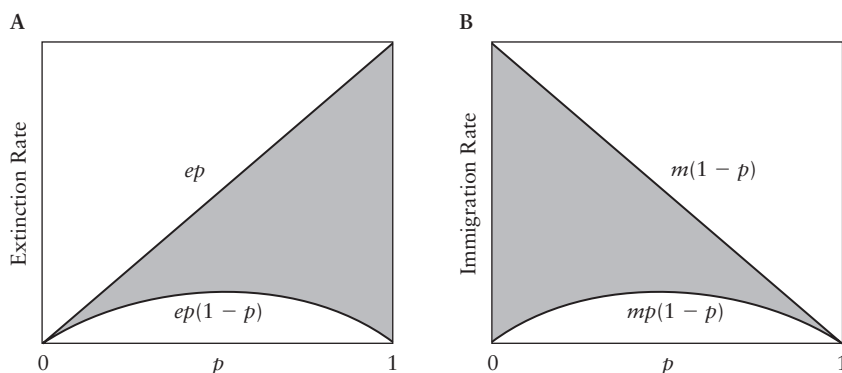


FIGURE 5.17. Differences in extinction rates (the rescue effect) and immigration rates (the propagule rain). (A) Extinction rates in the models of Levins (ep) (1969) and Hanski [$ep(1 - p)$] (1982); p is the fraction of sites occupied. The difference between the two curves (shaded region) constitutes the rescue effect. (B) Immigration rates in the models of Levins and Hanski [$mp(1 - p)$] and that of Gotelli [$m(1 - p)$] (1991). The difference between these two curves constitutes the propagule rain.

APPENDIX: DATA FOR EXERCISES 5.2, 5.3, AND 5.4

Species	x coord.	y coord.	Species	x coord.	y coord.
American Basswood	285	105	American Beech	209	351
American Basswood	293	132	American Beech	202	385
American Basswood	299	114	American Beech	278	114
American Basswood	277	83	American Beech	193	307
American Basswood	264	121	American Beech	230	256
American Basswood	300	127	American Beech	243	260
American Basswood	126	269	American Beech	251	271
American Basswood	290	99	American Beech	188	330
American Basswood	325	103	American Beech	191	327
American Basswood	234	260	American Beech	201	291
American Basswood	185	183	American Beech	196	329
American Basswood	197	248	American Beech	230	324
American Basswood	189	195	American Beech	171	319
American Basswood	327	86	American Beech	168	319
American Basswood	313	126	American Beech	252	270
American Basswood	294	98	American Beech	230	285
American Basswood	260	100	American Beech	346	138
American Basswood	317	110	American Beech	266	263
American Basswood	301	92	American Beech	178	338
American Basswood	208	170	American Beech	182	322
American Basswood	331	93	American Beech	201	258
American Basswood	173	226	American Beech	96	196
American Basswood	177	207	American Beech	264	286
American Basswood	199	187	American Beech	193	329
American Basswood	182	198	American Beech	344	144
American Basswood	286	104	American Beech	227	342
American Basswood	236	192	American Beech	71	54
American Basswood	324	104	American Beech	199	297
American Basswood	303	98	American Beech	187	329
American Basswood	305	92	American Beech	165	318
American Basswood	254	223	American Beech	160	324
American Basswood	301	94	American Beech	167	326
American Basswood	306	44	American Beech	220	247
American Basswood	219	185	American Beech	188	384
American Basswood	272	113	American Beech	176	332
American Basswood	297	132	American Beech	171	391
American Basswood	315	136	American Beech	235	367
American Basswood	301	90	American Beech	178	357
American Basswood	276	98	American Beech	216	109
American Basswood	338	73	American Beech	257	332
American Basswood	271	95	American Beech	269	67
American Basswood	313	125	American Beech	236	204
American Basswood	281	79	American Beech	190	342
American Basswood	279	133	American Beech	263	214
American Basswood	299	94	American Beech	188	250
American Basswood	292	102	American Beech	201	83
American Basswood	332	58	American Beech	238	312

Species	x coord.	y coord.	Species	x coord.	y coord.
American Basswood	309	128	American Beech	122	320
American Basswood	209	176	American Beech	268	211
American Basswood	324	57	American Beech	253	251
American Basswood	218	227	American Beech	209	217
American Basswood	332	53	American Beech	236	204
American Basswood	233	225	American Beech	269	67
American Basswood	301	145	American Beech	265	214
American Basswood	262	97	American Beech	230	186
American Basswood	295	96	American Beech	236	200
American Basswood	186	233	American Beech	269	67
American Basswood	194	230	American Beech	48	12
American Basswood	213	239	American Beech	64	13
American Basswood	188	233	American Beech	189	321
American Basswood	320	127			
American Basswood	281	98			
American Basswood	327	101			
American Basswood	315	79			
American Basswood	291	89			
American Basswood	189	229			
American Basswood	241	193			
American Basswood	305	65			
American Basswood	312	96			
American Basswood	313	96			
American Basswood	194	215			
American Basswood	296	114			

The first five chapters of this book have dealt with the situation in which all interactions are among individuals within a single species. Individuals interact in order to reproduce, thus creating a birth rate. They interact indirectly when they use the same resources, thus creating the phenomenon of competition. They interact in complicated ways to create nonlinear effects, especially in structured populations. This somewhat extensive introduction to population ecology examined many of these sorts of interactions but within the context of individuals interacting with one another in a single population or that of subpopulations interacting with one another in a larger population. Yet in most real ecosystems, critical interactions occur among species. A fir forest in northern Finland may be an excellent example of a population in which the dominant interactions are individual to individual within a single population of one species. But in a tropical rain forest in Borneo, hundreds of species of trees interact with one another in a multispecies context. There it is a case of populations interacting with populations in the context of competition among species of trees, which is the critical force structuring the ecosystem. Even the simple density-dependent models we examined were implicitly cases of populations interacting with one another. When two individuals of the same species consume the same resource, if that resource is actually another organism, as it frequently is, the critical interaction is between the population of the consuming species and the population of the species that is consumed—a consumer–resource or predator–prey interaction.

Depending on one's intentions in developing a model, it is frequently useful and certainly justifiable to take a phenomenological approach and simply model a single population as having density dependence without stipulating what is causing that effect. However, it is frequently the case that one needs to explicitly acknowledge the second species involved in the interaction and

model the system accordingly. Thus population ecology normally also includes the subject of populations of different species interacting with one another.

Although the classification of these interactions can become rather extensive and cumbersome (predator–prey, competition, mutualism, symbiosis, consumer–resource, herbivore–plant, facilitation, commensalism, and so on), from the point of view of developing basic theory there are really only three basic forms. First is the interaction that occurs when one species has a positive effect on a second species while the second species has a negative effect on the first (predator–prey, consumer–resource, herbivore–plant, parasite–host, facilitation–competition). Second is the interaction that occurs when both species have negative effects on one another (interspecific interference, interspecific competition). Third is the interaction that occurs when both species have positive effects on one another. We focus on each of these in the final four chapters before the conclusion, separately analyzing relatively large things eating relatively small things (classical predator–prey theory, in this chapter), very small things eating large things (classical epidemiology theory or disease ecology, in chapter 7), two similar populations striving to eat similar things (competition theory, in chapter 8), and two populations affecting one another positively (mutualism theory, in chapter 9).

EXERCISES

6.1 Suppose that two species are each growing according to the exponential equation

$$\frac{dN_1}{dt} = rN_1,$$

$$\frac{dN_2}{dt} = rN_2,$$

and remember that $r = b - m$, where b is birth rate and m is mortality rate. Further suppose that the first species is a predator that eats the second species, which means that its birth rate is proportional to the second species and the mortality rate of the second species is proportional to the second species. Write the resulting equations.

6.2 From the equations you derived in exercise 6.1, calculate the equilibrium conditions for both equations, and plot those conditions on a graph of N_1 versus N_2 , indicating the values of the intercepts. Note that each of the resulting lines is called an isocline.

Predator–Prey Interactions: First Principles

The elementary process of consumption, one organism feeding off another organism, was the main subject of the theoretical studies of Lotka (1926) and Volterra (1926), the influence of which cannot be overestimated. We begin with the simplest form of those theoretical studies.

Consider a prey, the population density of which is V (for victim), and a predator, the population density of which is P . Beginning with the basic rules of exponential growth, we could write

$$\frac{dV}{dt} = (b_1 - m_1)V \quad (1a)$$

and

$$\frac{dP}{dt} = (b_2 - m_2)P, \quad (1b)$$

where b and m are birth rate and mortality rate, respectively (see chapter 1). However, by the nature of these two populations we can also say that the birth rate of the predator cannot be constant; it must depend on the population density of the prey (we assume that there are no alternative food sources for the predator, so if the prey population is zero, the predator's birth rate must also be zero). Furthermore, the mortality rate of the prey must depend on the population density of the predator (we assume that there are no alternative sources of mortality other than being eaten by the predator). Assuming the simplest relations possible (as a first approximation), the birth rate of the predator can be cast as a positive linear function of the prey, or

$$b_2 = bV, \quad (2a)$$

and the mortality rate of the prey can similarly be regarded as a simple positive linear function of the density of the predator, or

$$m_1 = aP. \quad (2b)$$

Substituting equations 2a and 2b into 1a and 1b, we obtain

$$\frac{dV}{dt} = b_1V - aPV$$

and

$$\frac{dP}{dt} = bVP - m_2P.$$

Finally, to make this correspond to the notation normally found in the literature, we let $b_1 = r$ and $m_2 = m$, giving us

$$\frac{dV}{dt} = rV - aPV \quad (3a)$$

and

$$\frac{dP}{dt} = bVP - mP. \quad (3b)$$

Equations 3a and 3b are the classic predator-prey equations of Lotka and Volterra. We will return to them after taking a rather different approach to the elementary theory of predator and prey interactions. Rather than thinking of equations, we ask about the general behavior of a predator and prey when they interact with one another and depict these behaviors graphically. Let us presume that the predator will be able to increase its population only when there are a critical number of prey around. That is, if the prey population goes below

some critical number, the predator population cannot maintain itself and begins declining. Similarly, we can postulate a critical number of predators below which the prey population can “escape” predation to some extent and increase its population but above which the predator simply overwhelms it, causing it to decrease. These two critical points are illustrated in figure 6.1. Because these critical points separate the part of the space in which the population is increasing from that in which it is decreasing, the point itself must represent a population that is neither increasing nor decreasing. That point is referred to as an isocline. The isocline, formally, is the set of all points for which the derivative (with respect to time) is exactly zero, which is to say the set of points for which the tendency of the population to increase is exactly balanced by the tendency to decrease. We can put the two isoclines together and easily see the behavior of the two populations when they are interacting, as in figure 6.2.

In figure 6.2A the two isoclines separate four regions of the predator–prey space, one in which the predator increases and the prey decreases (quadrant I), one in which both predator and prey decrease (quadrant II), one in which the predator decreases and the prey increases (quadrant III), and one in which both predator and prey increase (quadrant IV). In figure 6.2B, two illustrative trajectories deduced from the simple vector field of figure 6.2A are presented. The most important conclusion to be drawn from this simple graphic analysis is that predator and prey will be oscillatory with respect to one another. This fact is intuitively obvious, at least with the help of the graph in figure 6.2A, and was the key insight of Lotka’s and Volterra’s analyses.

Although the above derivation of the isoclines is intuitive and certainly adequate, it is also possible to make the same derivation somewhat more formally. Returning to the formal model (equations 3a and 3b), we find the isoclines by setting the derivatives equal to zero, in which case we obtain

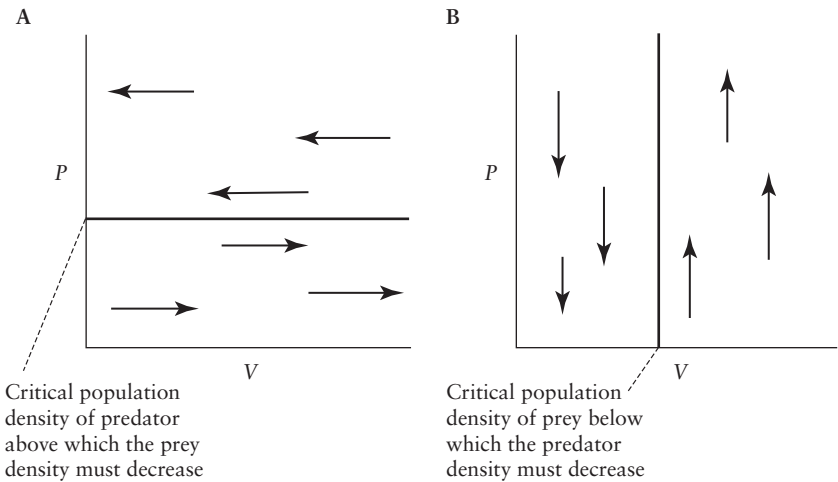


FIGURE 6.1. Qualitative construction of isoclines separating regions of space where (A) prey increases or decreases and (B) predator either increases or decreases. P , predator; V , prey (victim).

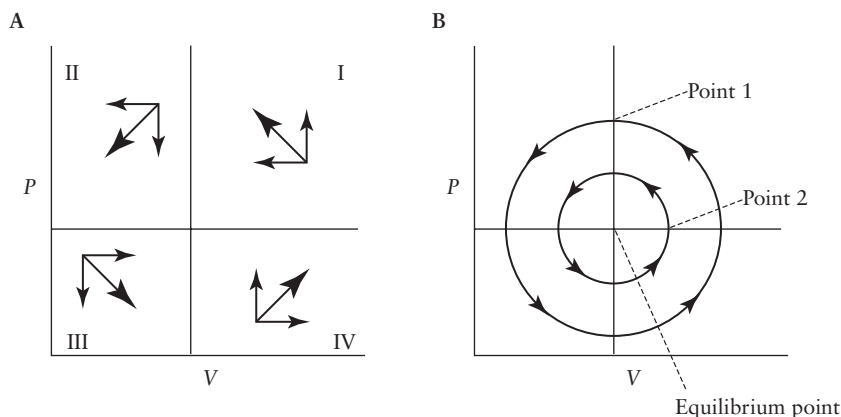


FIGURE 6.2. Construction of trajectories from the combination of the prey and predator isoclines in figure 6.1. (A) The basic isocline structure leads to the vector field (large arrows). (B) Two illustrative trajectories. (For an explanation of the quadrants in panel A and the points in panel B, see text.)

$$P = r/a \quad (4a)$$

for the prey isocline and

$$V = m/b \quad (4b)$$

for the predator isocline. These correspond to the intuitive “critical population density of prey below which the predator must decrease” and “critical population density of predators above which the prey must decrease,” as we defined these isoclines earlier in formulating figure 6.1. These isocline equations may initially seem confusing because, for example, the prey isocline is defined by the equilibrium density of the predator (P^*), whereas the predator isocline is defined by the equilibrium density of the prey (V^*). This situation can be made more intuitive by noting that the prey isocline is defined solely by a critical predator density and is independent of the prey density itself. Similarly, the predator isocline depends only on having enough prey around and is independent of its own density.

The two-species equilibrium point is where the isoclines intercept, that is, where both species are at equilibrium (see figure 6.2B). This equilibrium point is neither an attractor nor a repeller but is of neutral stability. Recalling the development in chapter 4, this situation is actually a bifurcation point and is structurally unstable, as will be apparent in a later section.

The most important prediction from the elementary theory is that predators and prey should oscillate with respect to one another, alternating between states of low predator/high prey numbers and low prey/high predator numbers. Does this prediction actually bear out in the natural world?

The most analyzed data set relating to this question is the lynx–hare data set accumulated over many years in Canada by the Hudson Bay Company. Based on the number of pelts brought in to the company each year and

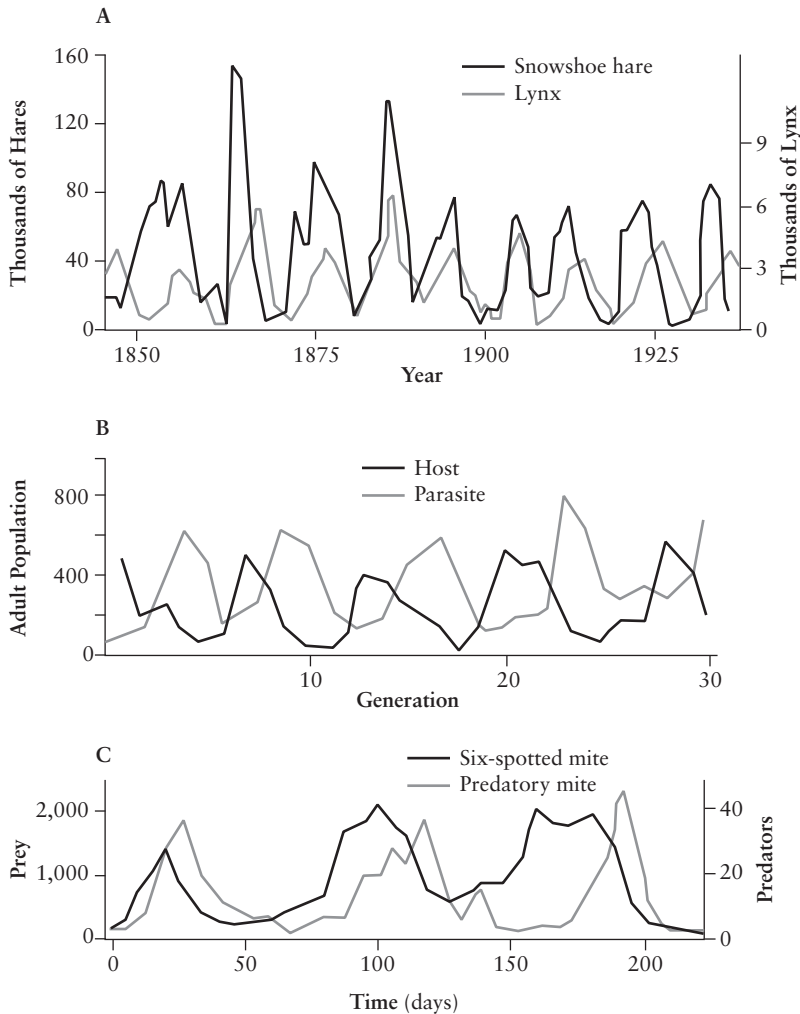


FIGURE 6.3. Examples of oscillating predator–prey dynamics. (A) Numbers of pelts of lynx and snowshoe hares brought in to the Hudson Bay Company by trappers. (B) Host and parasite numbers of the Azuki bean weevil (*Callosobruchus chinensis*) and its parasite (*Heterospilus* sp.) (Utida 1957). (C) Number of six-spotted mites (*Eotetranychus sexmaculatus*) and its predator (*Typhlodromus occidentalis*) in the experiments of Huffaker (1958).

assuming that the number of pelts that hunters and trappers were able to obtain was correlated with the actual population density of the two species, it is possible to estimate the numbers of individuals in the two populations over the general area traversed by hunters and trappers. These data have now been analyzed many times, frequently with great care and detail, revealing a variety of interesting details (e.g., Blasius et al. 1999). But the most obvious result can be obtained from a quick glance at a graph of the data (figure 6.3A). Predator and prey oscillate with respect to one another, just as the qualitative generalization of the theory shows.

Historically, laboratory studies have been perhaps even more convincing. For example, Japanese workers counted numbers of Azuki bean weevils and their parasites over a period of several months. The oscillations of predator and prey could not be clearer (figure 6.3B). In attempting to understand the dynamics of a parasite for the purposes of biological control, Huffaker (1958) set up a grid of oranges with complicated connections between them and followed the mite pest *Callosobruchus chinensis* and its predator *Heterospilus* over time. His results are displayed in figure 6.3C, again illustrating the qualitative conclusion that predator and prey tend to oscillate with respect to one another over time.

EXERCISES

- 6.3 Using the basic form of the Lotka–Volterra equations (3a and 3b), replace the exponential part of the “victim” equation with the logistic equation and solve for the equilibrium state, expressing the critical value of P as a function of V . Plot the isocline.
 - 6.4 Using the basic form of the Lotka–Volterra equations (3a and 3b), replace the exponential part of the predator equation with the logistic equation and solve for the equilibrium state, expressing the critical value of P as a function of V . Plot the isoclines.
 - 6.5 On the graphs you made in exercise 6.2, shade the four areas that represent the four different states possible: (1) predator increasing, prey increasing; (2) predator increasing, prey decreasing; (3) predator decreasing, prey increasing; and (4) predator decreasing, prey decreasing. Now keep the equilibrium point (the point where the two isoclines cross) the same but change the prey isocline so that it is declining. Remember that the isoclines still separate those four critical dynamic areas. Compare the general areas of the vector field (the collection of vectors that indicate the general dynamics of the system), and guess what the overall change in behavior of the system will be.
-

Density Dependence

The classic Lotka–Volterra predator–prey model leads to the important generalization, as discussed above, that predator and prey ought to oscillate with respect to one another. This makes some intuitive sense. However, it also results in a prediction that is universally agreed to be biologically absurd, that whatever the mix of predator and prey, the system will always return to exactly that mix sometime in the future—the case of neutral stability (technically referred to as centers). That states of neutral stability are useful points of reference in the global analysis of models was discussed in chapter 4, but here we seek to go beyond the simple assumptions of the Lotka–Volterra equations and add some measure of biological realism to the analysis of consumption, that is, predator–prey relations (which in more general terms are consumer–resource systems and include herbivores and plants, parasites and hosts, etc.).

Recalling the development of the logistic equation, we began with the process of exponential growth and proceeded by adding density dependence. Then we replaced the assumption that the per capita growth rate remains constant

with the assumption that it declines with population density. It would seem most obvious that we should do the same in the predator–prey situation. If we take the basic predator–prey equations already shown in equations 3a and 3b,

$$\frac{dV}{dt} = rV - aVP, \quad (5a)$$

and

$$\frac{dP}{dt} = bVP - mP \quad (5b)$$

using the same reasoning we used in deriving the logistic equation, it is not reasonable to assume that the per capita birth rate of the prey (equation 5b) is constant. This would be the same as presuming that the prey population could grow without limit if the predator was absent. This part of the equation can be made more biologically realistic by simply making the prey population obey the logistic equation when the predator is absent. So equations 5a and 5b become

$$\frac{dV}{dt} = rV \left(1 - \frac{V}{K}\right) - aVP, \quad (6a)$$

and

$$\frac{dP}{dt} = bVP - mP \quad (6b)$$

where K is the carrying capacity of the prey population. The isocline for the prey equation (equation 6a) is

$$P = \frac{r}{a} - \frac{r}{aK}V,$$

which is a linear equation in the P, V phase plane. The qualitative dynamics that result from adding density dependence to the equation for V are deducible from an examination of the isoclines and how they change when density dependence is added, as illustrated in figure 6.4. In the upper graph of figure 6.4, the original isocline is indicated by a dashed line and the original vectors are also dashed.

The new isocline (the isocline that exists after density dependence is added to the prey equation) changes a bit of the region that used to be part of quadrant II or IV into part of what is effectively quadrant III or I (respectively). The change in vectors in this changed bit of those quadrants is shown, and from this it can be seen that a bit of the space now has vectors that point toward the equilibrium point more than they did with the original isocline. This causes the oscillations to move toward that equilibrium point. Thus adding density dependence to the prey population causes the neutral oscillations to change into stable oscillations, which is to say that the system becomes an oscillatory point attractor. This is not a trivial conclusion, because it restricts the behavior of predator–prey systems considerably, probably also unrealistically. It suggests that predator–prey systems are always stable! This is also an unwarranted conclusion, as discussed in the following section.

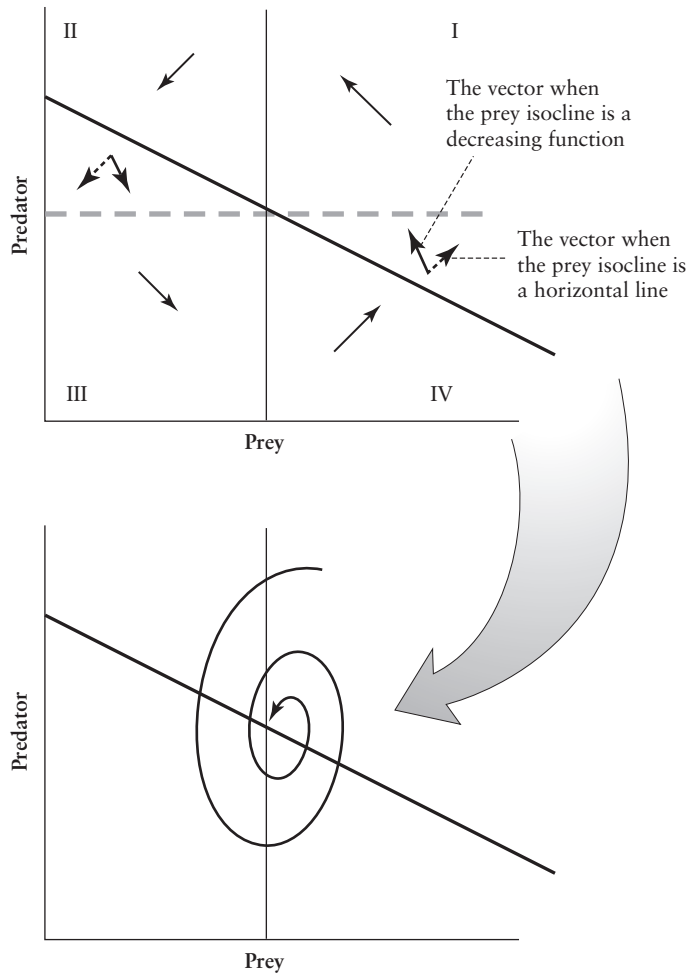


FIGURE 6.4. The changes in the isoclines and population trajectory from adding density dependence to V (see text). Dashed lines are the original vectors and prey isocline.

EXERCISES

- 6.6** Equations 6a and 6b can be thought of as each having two parts, a growth part and a mortality part. For the prey equation, the growth part is

$$rV\left(1 - \frac{V}{K}\right),$$

and the mortality part is (aVP) . The mortality part says how the prey population is being reduced by the predator, which is basically the same as how many prey the predators are eating, or the predation rate. So each predator eats at the rate aV . Make a graph of the predation rate as a function of V , assuming that $a = 0.5$ (i.e., during the relevant time period, each predator eats $0.5V$ prey). Examining the graph, does

it make biological sense? Imagine yourself as a predator eating olives. If there are 2 olives on your plate, how many of them will you eat? If there are 10, how many will you eat? If there are 10,000, how many will you eat? Sketch onto the graph what you think might be a more biologically reasonable assumption about the shape of that curve (ask yourself how many of the 10,000 olives you are really likely to eat).

- 6.7 Carnivorous tunicates are sessile, so they must wait for their food to come their way. How many food items will a tunicate eat in a given period of time? Suppose that the overall time they allot to eating during the day is 12 hours (tunicates sleep the rest of the time), and each time a food item is taken in, an individual tunicate requires half an hour to swallow it. The frequency at which food items come its way will depend on the density of those food items. The time required to encounter a food item will be inversely related to the density of the food items, or $1/V$. What is the equation that relates the total number of food items eaten in a day to the population density of the food? Make a graph of the equation for values of V from 0 to 30.
- 6.8 If the number of prey eaten is V_e and the number available is V , rewrite the equation derived in exercise 6.7 after transforming the variables to $v_e = 1/V_e$ and $v = 1/V$ such that v_e is presented as a function of v . What can you say in words about the relationship between the inverse of food eaten as related to the inverse of food offered?
- 6.9 The following pairs of numbers represent the number of aphids eaten and the number offered in an experiment with a ladybird beetle and its aphid prey: 18, 25; 34, 50; 60, 100; 95, 200; 108, 300; 122, 400; 125, 500; 132, 600; 141, 700; and 145, 800. Use the equation you derived in exercise 6.8 to estimate a general function that fits the data.

Functional Response

The density-independent assumption for the prey species seems quite foolish biologically. However, the parallel assumption for the predator is not really all that unreasonable, because competition for resources is built into the model indirectly—the amount of food available already depends on the density of predators as defined in the equation for prey density). The equation simply says that the predator will have a positive birth rate as long as there is food to eat, which is sensible under the restrictive set of assumptions used in developing these models. However, another assumption of the model about predator behavior does seem quite unreasonable. We have presumed that the ability of the predator to eat its prey is totally independent of the density of the prey, and similarly that the effect of the predator on the prey population is independent of the density of the prey population. These are the assumptions that we make in asserting that r and a are constants.

It is actually well documented (e.g., Hassell 1978) that prey are not eaten independently of the prey density. Indeed, for most predators, if you plot the rate of prey consumption against the population density of the prey, you do not get a straight line, as would be predicted by equations 6a and 6b (i.e., the predation rate per predator is aV (the overall rate is then aVP),

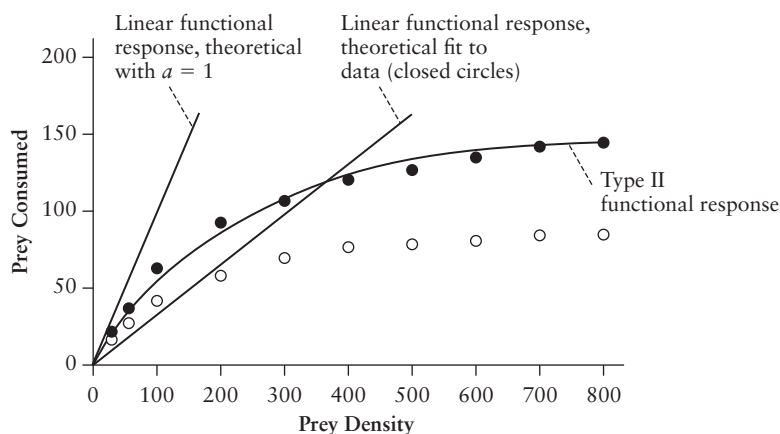


FIGURE 6.5. Type II functional response of ladybird beetles *Coccinella transversalis* (closed circles) and *Propylea dissecta* (open circles) at different densities of the aphid *Myzus persicae* in India (data from Pervez 2005).

which means that the predation rate is a linear function of the population density of the prey). For example, in figure 6.5 we see the results of an experiment with the ladybird beetles *Coccinella transversalis* and *Propylea dissecta* preying on the aphid *Myzus persicae*. The original Lotka–Volterra equations assume that predators will eat a certain fraction of prey no matter how many are available. Thus if a lion eats 1% of the zebra population each month, that means that it will eat 1 of a population of 100, 10 of a population of 1,000, and 10,000 of a population of 1,000,000. That is, it will never become satiated. This is clearly an unreasonable assumption for most predators.

Yet that is exactly the assumption made by the classic form of the Lotka–Volterra equations. In terms of actual data, the assumption is obviously ridiculous. For example, we have plotted the Lotka–Volterra assumption as a straight line with intercept at zero (that is, after all, what the assumption says) for the data in figure 6.5. In fact what we actually see in almost all experiments is that the rate of predation is a curvilinear function of the population density, much like the function labeled “Type II functional response” in figure 6.5, and a linear approximation is clearly unwarranted.

The predation rate (the number of prey consumed per predator) is always a function of the population density of the prey, and that is why it is usually referred to as the functional response. However, the exact form it takes can have a dramatic effect on the qualitative outcome of the predator–prey interaction. As we have already seen, if the functional response is linear, the system either is one of neutral stability (if the prey are not density dependent) or is an oscillatory point attractor (if the prey are density dependent). What if the functional response is nonlinear?

This form of the function is most frequently modeled with the so-called Holling disk equation, after Holling (1959), which states

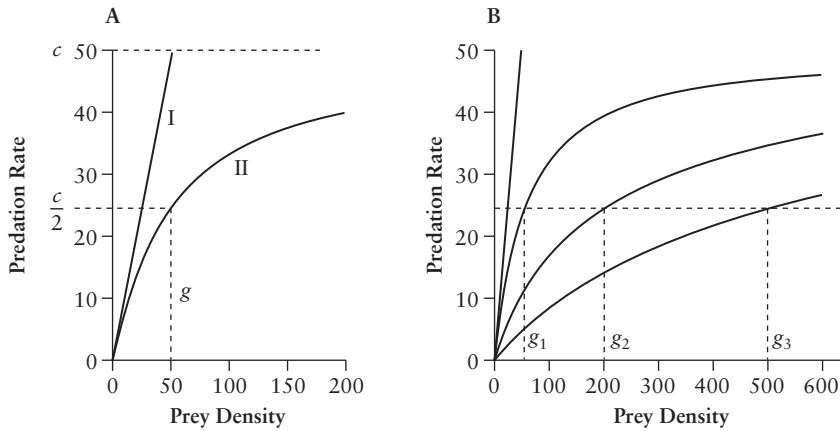


FIGURE 6.6. Forms of the functional response. (A) Linear form (I) and asymptotic form (II). (B) Three versions of a type II functional response with different values of the parameter g ($g = 50, 200$, and 500). The parameter $c = 50$.

$$\text{Predation rate} = \frac{cVP}{g + V},$$

an example of which is shown in figure 6.6A labeled II. The meaning of the parameters in the Holling disk equation are illustrated in the figure. The value c is the asymptotic (maximum) value of the predation rate per predator, and g is the value that V assumes when the predation is at half its maximum value. Thus c is the asymptotic value, whereas g is a measure of the degree of curvature of the curve. As g increases, the curvature of the functional response slowly increases. Changes in the form of the functional response are illustrated in figure 6.6.B.

In addition to the strict mathematical meaning of the parameters c and g (which are made clear in figure 6.6), it is possible to ascribe more biological meaning to the parameters by deriving this form of the functional response from first principles. Every predator needs to spend time finding prey. The time it takes to encounter an individual prey will be an inverse function of the prey in the system (i.e., time to encounter a prey = a/V), such that when there are large numbers of prey, the time it takes to encounter one of them is very short, but when there are very few prey, the time it takes to encounter one of them is long. After encountering prey, the predator needs to process it (kill, ingest, and digest it), commonly referred to as handling time. So the total time needed for a predator to deal with a single prey will be $(a/V) + b$, where b is the handling time. The rate at which prey will be found and handled will then be the inverse of the time it takes to deal with an individual prey, which is to say $1/[a/V + b]$, which we can rearrange to be $(1/b)V/[(a/b) + V]$. Thus we see that the parameter c is the inverse of the handling time, representing the maximum possible number of prey that a predator could eat if prey were instantly available, that is, if there was no encounter time and consumption was limited

only by how long it took to process each prey. The parameter g is the ratio of the encounter time to the handling time; the functional response is closer to linear the less important handling time is relative to encounter time.

To incorporate a nonlinear functional response into the prey equation, we start with its verbal description, $dV/dt = \text{rate of growth in the absence of a predator} - \text{predation rate}$, which is

$$\frac{dV}{dt} = rV - aVP$$

when the predation rate is a linear function of the prey density. The rate of growth in the absence of a predator is rV , and the predation rate is aVP . But this is assuming that the prey population is density independent (see previous section) and that the predation rate is linear. If the predation rate is nonlinear, we can employ the disk equation of Holling to obtain

$$\text{Predation rate} = \frac{cVP}{g + V},$$

which makes the prey equation

$$\frac{dV}{dt} = rV - \frac{cVP}{g + V},$$

which has the isocline

$$P = \frac{r}{c}(g + V).$$

EXERCISES

- 6.10** Plot the type II functional response for $c = 100$ and $g = 100, 300$, and $1,000$.
- 6.11** The prey isocline with the type II functional response added is $P = (r/c)(g + V)$. Graph that for $r = 1.4$, $g = 50$, and $c = 50$. Graph it again for $r = 1.5$, $g = 100$, and $c = 95$.
-

Following the above argument, we must also modify the predator equation. That is, if the predation rate is an inverse function of the prey density, that fact must also be represented in the predation rate of the predator equation. So, if the predator equation is

$$\frac{dP}{dt} = \text{birth rate} - \text{death rate}$$

and, as argued before, the birth rate is a function of the prey density; it is the birth rate that becomes satiated. That is, the birth rate cannot be a linear function of V , because the linear assumption implies that the predator can effectively eat an infinite amount of food. So the birth rate of the predator must be modified in the same way as the prey equation. Thus, if the original equation was

$$\frac{dP}{dt} = bVP - mP,$$

where bVP is the birth rate, we must modify that birth rate with the functional response. Thus the predator equation becomes

$$\frac{dP}{dt} = \frac{bVP}{g + V} - mP,$$

which has the isocline

$$V = \frac{mg}{b - m},$$

which is qualitatively the same as without the functional response term, that is, a simple vertical line.

The two isoclines are plotted in figure 6.7, along with a qualitative interpretation of the vector field. Note that here we have a point repeller. That is, with

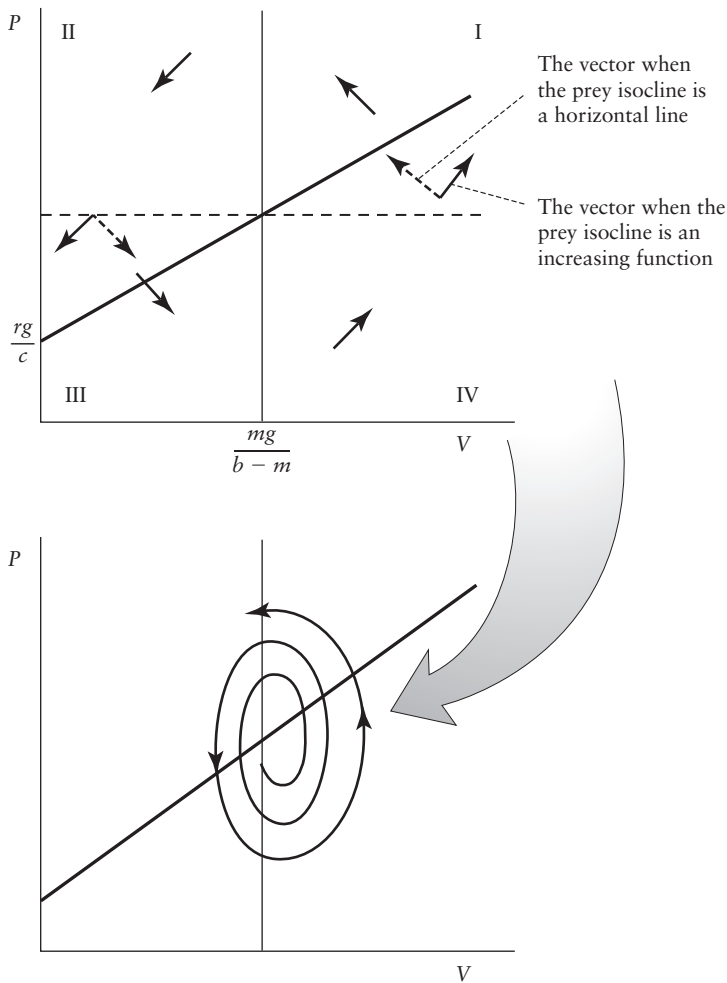


FIGURE 6.7. The changes in the isoclines and population trajectory from adding a type II functional response.

the simple addition of a nonlinear functional response to the classical Lotka–Volterra equations, we conclude that all predator–prey systems are unstable and thus cannot persist!

So from two separate simple and biologically sensible modifications of the classical Lotka–Volterra predator–prey equations we conclude, first, that all predator–prey systems are stable (have oscillatory attractors), and second, that all predator–prey systems are unstable (have oscillatory repellers).

EXERCISES

- 6.12** Combine the prey equation that includes density dependence (equation 6a) with the idea of a type II functional response. Write the differential equation, and solve for the isocline.
- 6.13** Using the equation derived in exercise 6.12, make a graph of P versus V for the following four parameter sets: $r = 5, c = 50, K = 100, g = 25$; $r = 5, c = 50, K = 120, g = 25$; $r = 5, c = 50, K = 100, g = 50$; and $r = 5, c = 50, K = 80, g = 25$.
-

Functional Response and Density Dependence Together

The truth of the matter is that the two nonlinear forces of functional response and density dependence seem to balance one another out. That is, the tendency of predator–prey systems to be unstable (resulting from the nonlinear functional response of the predator) is counterbalanced by their tendency to be stable (resulting from the density dependence of the prey). This can be seen if we simultaneously add both nonlinear components to the original equations. For the prey equation we have

$$\frac{dV}{dt} = rV \left(\frac{K - V}{K} \right) - \frac{cVP}{g + V}.$$

The isocline for this equation is a bit more complicated than before and is given as

$$0 = rV \left(\frac{K - V}{K} \right) - \frac{cVP}{g + V},$$

which, after some algebraic manipulation, becomes

$$P = \frac{r}{c} \left[g + \left(1 - \frac{g}{K} \right) V - \frac{1}{K} V^2 \right]. \quad (7)$$

The important thing about equation 7 is that it is quadratic in the space of P, V , which means that it is shaped like a parabola, as illustrated in figure 6.8A.

From a simple qualitative point of view, the key feature of equation 7 is that the possibility exists for the prey isocline to cross the predator isocline

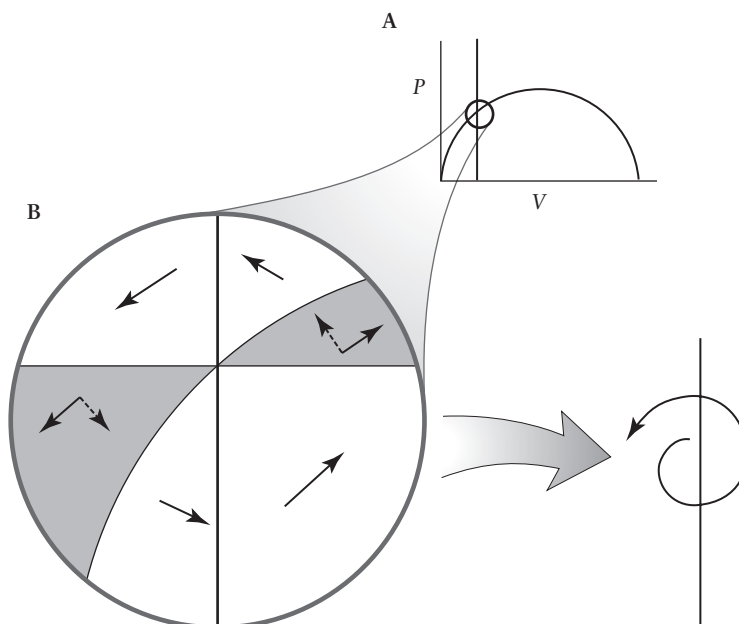


FIGURE 6.8. Changing dynamics due to adding both density dependence and a non-linear functional response to the Lotka-Volterra predator-prey equations.

when it is either ascending (as in figure 6.8A) or descending, depending the shape of the prey isocline or the placement of the vertical predator isocline. Looking at a close-up of the intersection when the predator isocline crosses to the left of the hump of the prey isocline (figure 6.8B), we see dynamic results, which are similar to the case of figure 6.7 except that the isocline is not strictly linear. The change in the prey isocline from a horizontal line to an ascending function has caused the dynamics to destabilize, and the intersection point is actually a repeller. The predator and prey have expanding oscillations, coming ever closer to the origin, which is to say ever closer to the extinction of the predator from the system. Because of the precise way in which the predator-prey equations are formulated, it is actually not possible for the predator to become extinct, but it is indeed possible for it to become “practically” extinct in that its numbers go so low that extinction through some random event is almost inevitable. As a rule of thumb, the closer the predator isocline is to the origin, the lower are the low points in the oscillations of the predator, as illustrated in figure 6.9.

On the other hand, recalling the previous graphical analysis in figure 6.4, one can easily see that if the predator isocline crosses the prey isocline to the right of the hump of the prey isocline when the prey isocline is descending, the oscillations will dampen (i.e., the equilibrium point will be a focal point attractor). But also note that this powerful rule of thumb works only if the predator isocline is a simple straight vertical line.

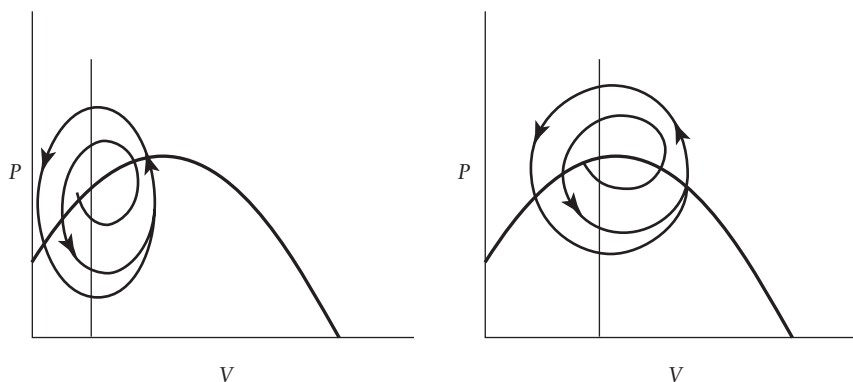


FIGURE 6.9. Trajectories of the predator–prey system with the predator isocline intersecting the prey isocline in its ascending limb.

EXERCISES

- 6.14** Repeat exercise 6.13 but with the parameters r , c , and g set at 1, 50, and 25, respectively, and vary K from 50 to 100 to 130 to 150. Plot the results on a graph of P versus V .
- 6.15** Repeat exercise 6.13 but with the parameters r , c , and K set at 1, 50, and 200, respectively, and vary g from 10 to 35 to 50 to 100.

Paradoxes in Applications of Predator–Prey Theory

Because of the inherent nonlinearities involved in predator–prey theory, there are sometimes unexpected consequences of particular actions that might be taken in management situations or that might result from evolutionary changes in parameters. Here we cite a couple of those so-called paradoxes.

First, consider the case of enrichment of the environment. This is something that is natural for a manager to try in a variety of situations. In a lake, for example, adding fertilizer would seem to be a good way to increase primary production, which would increase the carrying capacity of the zooplankton, which would, in turn, increase the catch of fish. Or so it would seem.

To examine this situation we assume, quite reasonably, that increases in productivity can be represented by increases in the carrying capacity of the prey, K . We then examine how the predator responds to increases in K through changes in the stability properties of the new intersection points of the predator and prey isoclines. Recall the prey isocline,

$$P = \frac{r}{c} \left[g + \left(1 - \frac{g}{K} \right) V - \frac{1}{K} V^2 \right].$$

Calculate the prey and predator densities at the peak of the hump of this function by first differentiating

$$\frac{dP}{dV} = \frac{r}{c} \left[\left(1 - \frac{g}{K} \right) - \frac{2V}{K} \right]$$

and then setting the derivative equal to zero and solving for V :

$$V = \frac{K - g}{2}.$$

This is the maximum density of prey that can be observed at a given carrying capacity. To determine the predator density at this peak, we substitute the equilibrium value of the prey density into the equation for its isocline, namely

$$P = \frac{r}{c} \left[g + \left(1 - \frac{g}{K} \right) \left(\frac{K - g}{2} \right) - \frac{1}{K} \left(\frac{K - g}{2} \right)^2 \right],$$

which simplifies to

$$P = \frac{r}{4c} \left(K + \frac{g^2}{K} + 2g \right).$$

These points are plotted for a range of prey carrying capacities to yield the range of prey isoclines in figure 6.10. Note that the predator isocline does not change, because it does not depend on the carrying capacity of the prey (see above).

In this example, before enrichment the system was a focal point attractor (i.e., a stable point), whereas after enough enrichment the predator isocline intercepts the prey isocline to the left of its peak, and thus the system is a focal point repeller (unstable). Hence the paradox that enriching the environment has caused the system to shift from stability to instability (Rosenzweig 1971).

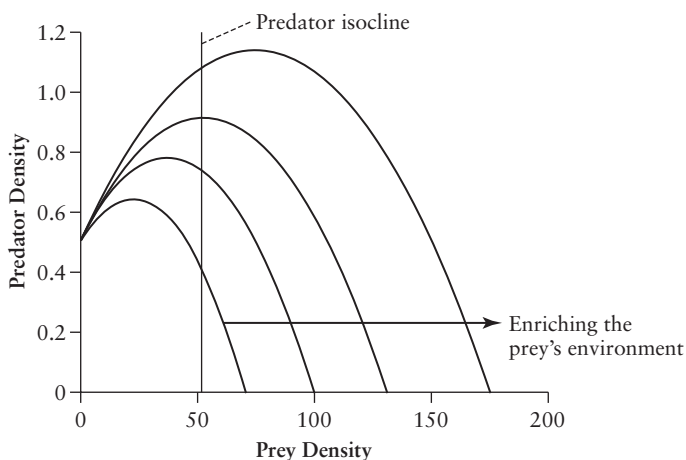


FIGURE 6.10. The paradox of enrichment. As the environment is enriched, the K of the prey species is increased. However, the predator isocline remains constant. The result is that, although the predator does increase its average density, it enters into an oscillatory mode that may in the end lead to its extinction. (Parameters for the prey isocline in this example are $r = 1$, $c = 50$, $g = 25$, and K varies from 70 to 100 to 130 to 175.)

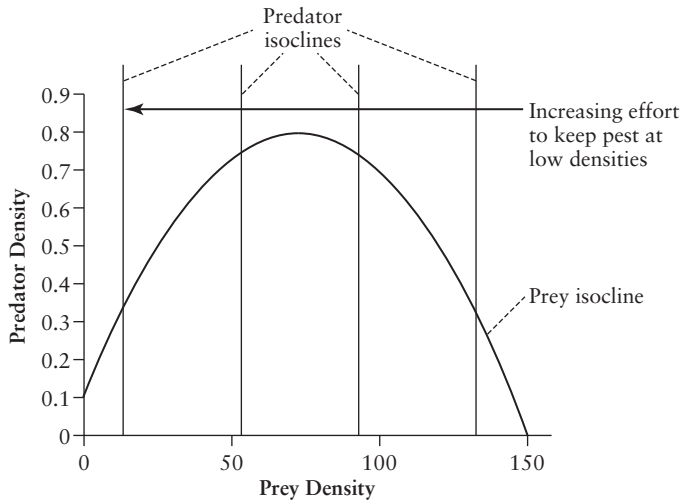


FIGURE 6.11. Isocline arrangements illustrating the paradox of biological control, i.e., that the two goals of such control are clearly incompatible in theory. As the pest manager moves the predator isocline to the left in an attempt to decrease the prey population, although the equilibrium value of the prey decreases, eventually the system becomes oscillatory, generating potentially damaging pest outbreaks.

Concrete examples of the paradox of enrichment are not readily available in the literature, although there are some suggestive examples. For example, populations of the larch bud moth in Switzerland seem to exhibit cycles in the optimal part of its altitudinal range but are relatively constant in numbers in marginal (less “enriched”) habitats (Baltensweiler 1971), which could be an example of the paradox (May 1981).

The paradox of biological control is another example. Biological control is effectively the application of predator–prey theory to the general problem of controlling pests in agriculture and forestry (Hawkins and Cornell 1999). Clearly, what is desired is that the predator (the biological control agent) maintain the pest at as low a density as possible. Yet the elementary theory of predator–prey dynamics shows quite clearly that if the two-species equilibrium point is at a low value for the prey, it is more likely to be unstable (see figure 6.9), and it is likely that the biological control agent will be eliminated from the system (Arditi and Berryman 1991). So if the desire is to develop a system that will maintain itself indefinitely yet cause the prey to be present at very low levels, we seem to have a paradox (figure 6.11). In a later section we discuss how this paradox was resolved in an actual case of biological control.

Predator–Prey Dynamics: A Graphical Approach

The above development is based on an analysis of equations—the predator–prey equations with density dependence added to the prey and with the

predator having a functional response of type II. It is useful to demonstrate how the same analysis can be heuristically presented in a simple graphical form following the pioneering work of Rosenzweig and MacArthur (1963) and Noy-Meir (1975). Begin by plotting the rate of change of the prey population (dV/dt) as a function of its own density (V). Note that this is *not* an isocline, where the densities of two species are plotted against each other, but rather a plot of the rate of change of one species versus the density of that same species (part of the idea is to derive the isocline from simpler first principles). At a very low density the rate of change of the total population will be small because there are so few individuals in the population. As the population density increases, so will the rate of increase, because more individuals will be reproducing. But we also expect that at high population densities the density-dependent effect will become important and the rate of increase will begin to decline. Finally, at very high population densities, near the carrying capacity of the environment, the rate of population increase will approach zero. In summary, we can expect, qualitatively, (1) an increasing population growth rate as population density increases from small values but (2) a decreasing population density as population density increases from larger values, with (3) zero population growth at both zero population density and carrying capacity. The resulting curve is indicated in figure 6.12, where it is labeled “Prey growth response.” Again, remember that the curve in figure 6.12 is not an isocline but the result of plotting the rate of growth against the population density, not the population density of one species against that of another species.

We can also plot the functional response of a predator population on the same axes because the functional response of the predator to prey density is actually a decrease of the prey population, although it is plotted in figure 6.12 as positive so we can compare it to the prey growth response. (Note

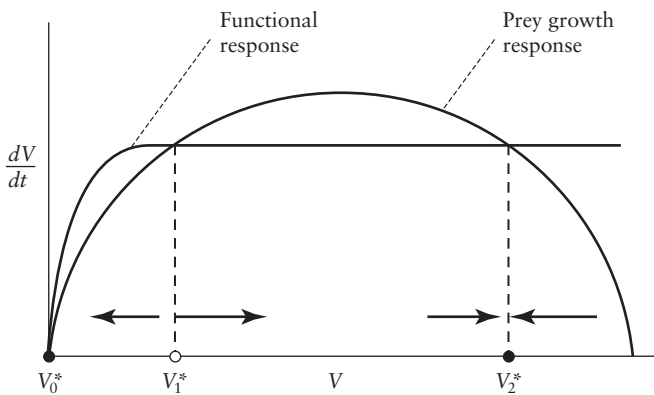


FIGURE 6.12. The two dynamic functions that make up the prey dynamic system, where functional response represents the mortality rate of the prey due to the predator.

that the functional response described earlier and shown in figure 6.6 is typically considered on a per-predator basis; here we plot the total amount consumed by a predator *population*, assuming that the predator population is constant and not dynamically responding to the prey population.) As we argued before, with a small number of prey, even a single predator will eat all or most of them. With more prey the predator will eat more of them but will eventually become satiated. For example, let us suppose that there are a fixed number of predators in the environment, say P^* . For that density of predators there will be a characteristic rate of depletion of prey, depending on the number of prey in the environment. With a small number of prey we might expect all of them to be eaten. If the predators were not satiated, we could add yet more prey to the environment and they would all, or almost all, be eaten. But ultimately we must be able to add so many prey that all of the predators in the environment will be satiated. After that point, even if we supersaturate the environment with prey, the rate of consumption of those prey will remain constant because the predators are satiated. Thus we expect that a graph of predation rate versus prey density will be an ascending curve with diminishing returns, as illustrated in figure 6.12, where the curve is labeled “Functional response.”

Examining figure 6.12, we can compare the growth rate of the prey population with the functional response (effectively, the mortality rate of the prey) for particular values of V . If the growth rate is greater than the functional response, the overall population will increase (because the growth rate is the number of prey increasing per unit of time and the functional response is the number of prey decreasing per unit of time), whereas it will decrease whenever the functional response is greater than the growth rate. Thus, between the two points V_1^* and V_2^* in figure 6.12, the prey population will be increasing, as indicated by the arrows pointing to the right on the graph. However, to the right of V_2^* , the functional response is greater than the growth rate, and we can expect the population to decline. Also to the left of V_1^* , the growth rate of the prey is less than the functional response, and we can expect the population to decline. The population declines are indicated by arrows pointing to the left in figure 6.12. A glance at this graph reveals the expected dynamics. V_2^* is an attractor, as is V_0^* , whereas V_1^* is a repeller. The point V_1^* is also sometimes referred to as a break point (May 1977) because it represents the value of the prey population density that can “break” a declining population and turn it into an increasing population.

This particular formulation is from Noy-Meir (1975) and was originally devised in the context of grazing mammals, in which the predators are herbivores and the resources are plants. The formulation intuitively makes clear why the prey isocline has the shape it does. If we plot functional responses for different numbers of predators in the system, we expect the functional response to increase as the number of predators in the system increases, as illustrated in figure 6.13. This means that the exact location of the attractor and repeller (the intersection of the functional response curve with the growth rate curve for prey) will change accordingly, as illustrated in figure 6.13. Plot-

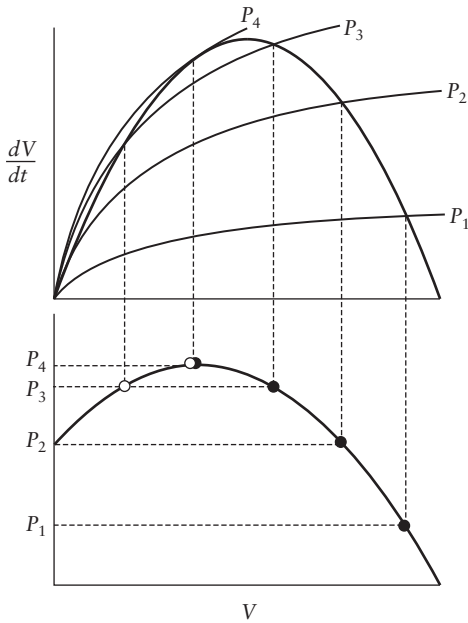


FIGURE 6.13. Constructing the prey isocline from the attractors and repellers of the graph of rate versus density. The upper graph shows the rate of growth of prey plotted along with four type II functional response curves. Where the rate of growth intersects the functional response is an equilibrium situation, which is then translated to the lower graph for those four particular predator densities. Stable points are indicated with solid circles and unstable points with open circles, and the points are connected with a dashed curve that represents the prey isocline.

ting the attractors and repellers on a graph of V versus P , we obtain the points plotted on the lower graph in figure 6.13B. It is clear that the prey isocline (where the prey is neither increasing or decreasing, as defined before) appears with a hump on a graph of predator versus prey. Note that this curve is the same shape as the one derived analytically in equation 7, and in fact the argument is actually quite the same, made more mathematically in leading up to equation 7 and more graphically here.

The basic shape of the prey isocline can also be deduced intuitively from reasoning about the isocline itself, as was done by Rosenzweig and MacArthur (1963). In the absence of a predator population, the prey population is expected to have some lower value below which it is unable to survive. For example, in the case of sexual animals, there is some low population density below which individuals of the opposite sex are unable to find one another very easily, and therefore the population will tend to decrease. At the other extreme, in a very dense population, the population will tend to decline because it is above its carrying capacity. Thus we expect the sort of dynamic behavior illustrated in figure 6.14A. The middle point (an open circle) is a repeller, and both the upper point (the carrying capacity) and the lower point (extinction) are attractors. Now let us presume that a small number of predators are added to the system but in such a way that they are nonreproductive and immortal—that is, we presume for temporary convenience that the predator population is unchanging. What changes are likely to occur in the dynamics of the prey population? First, it is likely that the upper attractor will be lowered due to the fact that the predators are eating a fixed number of prey during each time period. Second, it is likely that the lower repeller will be

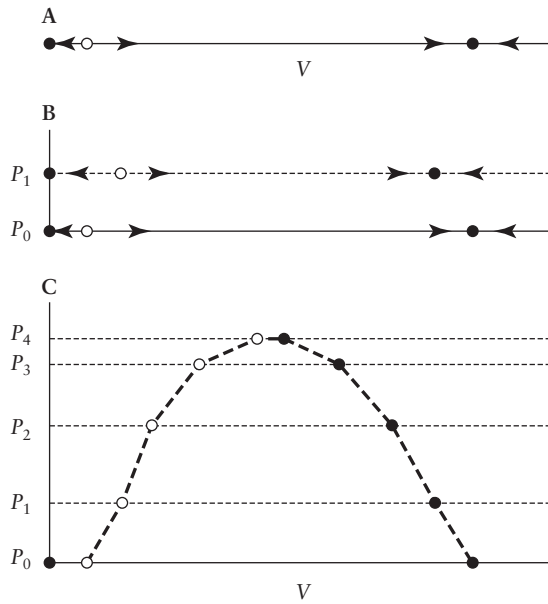


FIGURE 6.14. Construction of the prey isocline using the graphic arguments of Rosenzweig and MacArthur (1963).

higher, that is, that the point at which the prey population is unable to sustain itself—to maintain its birth rate higher than its death rate—will be a larger value. It will require a larger inoculum of prey organisms to form a successful population. This is illustrated in figure 6.14B. If we now repeat this thought experiment of adding a fixed predator population to the system, we can trace all the changes in both the repeller and the carrying capacity, as is done in figure 6.14C. It is evident that a hump-shaped curve results, corresponding to the analysis of Noy-Meir as presented above and the more mathematical analysis incorporating the density dependence of the prey and the type II functional response of the predator.

The Rosenzweig–MacArthur graphical approach has been used to model a variety of other qualitative situations, the most important of which is the addition of a refuge for the prey. If there is some part of the environment in which the prey population is invulnerable to the predator, there is a portion of the graph in which the prey population is capable of increasing even though the predator population may be very large. Graphically this is represented by a steep rise in the predator isocline at some low population density, specifically the population density that represents the size of the refuge (in terms of the number of prey it will protect). The resulting qualitative appearance of the prey isocline is illustrated in figure 6.15, along with the expected dynamic behavior. Clearly a refuge for the prey “stabilizes” the predator–prey oscillations (in the sense that the limit cycle is constrained at a higher prey value than when the refuge was not there).

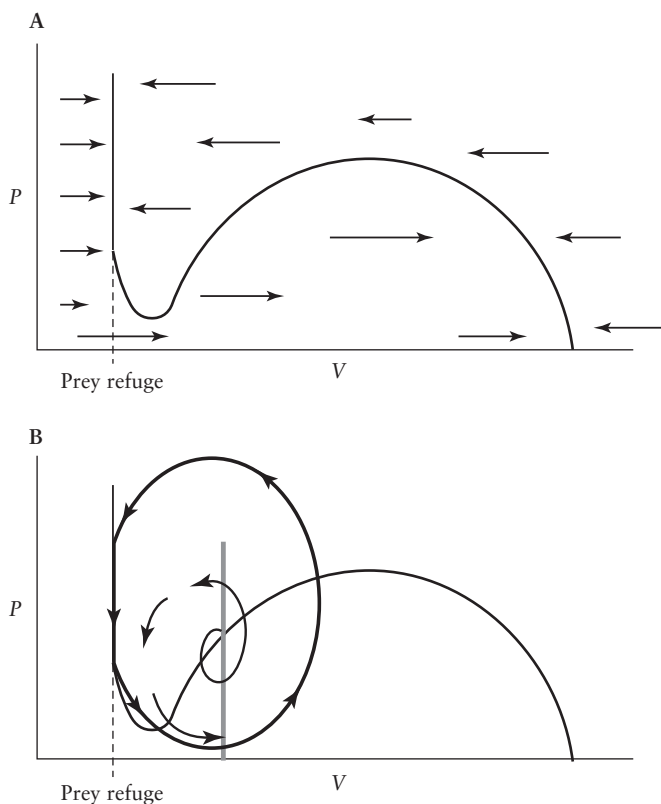


FIGURE 6.15. The prey isocline and its dynamics under conditions of a prey refuge. (A) The prey isocline. (B) The dynamic consequences.

A further complication is easily seen using the Noy-Meir approach. The type of functional response presented in figure 6.12 is probably the most common, but it is only one of three general types of response, as first noted by Holling (1959). The linear functional response simply proposes that the predator never becomes satiated. Although this may seem to be truly the case for some unwanted house guests, for the most part, consuming organisms get full, and they cannot ingest past some upper limit. Thus most ecologists agree that the functional response must have some upper limit. But it is also sometimes the case that predators are able to fully exploit prey only after they have learned of their presence or learned how to catch and eat them. And when prey are rare in the environment, there is less opportunity to learn of their presence as a potential food source than when they are common. Thus the rate of predation increases very slowly at low prey densities but accelerates very rapidly after the predators have learned to eat the prey. We thus expect a differently shaped curve for a “learning” predator, one that shows an accelerating rate of predation at very low prey densities, followed by a decelerating rate and ultimately, of course, by satiation. Thus the functional response curve should have a “sigmoid” shape, as indicated by the function labeled III in figure 6.16.

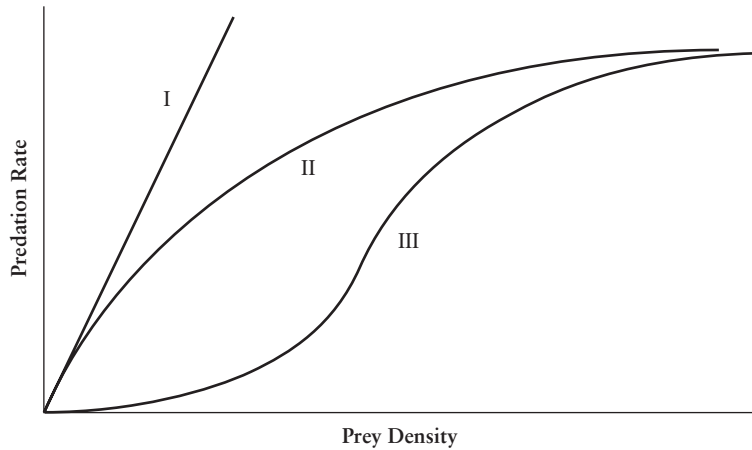


FIGURE 6.16. The three forms of the functional response.

Repeating the exercise that led to figure 6.13, the prey isocline is illustrated in figure 6.17 for the case of a type III functional response curve. Recalling the qualitative construction of the isocline with a prey refuge (see figure 6.15), we see that a type III functional response has an effect that is similar to that found with the addition of a prey refuge to the system (see figure 6.15).

The addition of a type III functional response, which in effect mimics the refuge addition in the graphical model, is a convenient way to theoretically resolve the paradox of biological control that was described earlier. If there is a refuge for the prey (or any of the other potential mechanisms that creates a prey isocline with a descending limb at low prey densities), the biological control paradox is resolved. The predator population can be retained in the system in a limit cycle with the prey, because the latter is somehow protected from the predator when at low densities. For example, in Murdoch's classic studies of the California olive scale, careful estimation of parameters suggested that the system was unstable. Yet the parasitic wasp that was the biological control agent appeared to persist in the environment, maintaining the scales at a very low density. Careful observation revealed that scales accumulated on the trunks of the olive trees where the parasites could not find them; the trunks effectively served as a refuge in which a small, low-density population of scale insects could survive regardless of the population density of the parasites (Murdoch et al. 1996).

Predator–Prey Interactions in Discrete Time

An alternative approach to predator–prey interactions is the well-known model of Nicholson and Bailey (1935), which is effectively a discrete form of the predator–prey equations of Lotka (1926) and Volterra (1926). Recalling the presentation of the one-dimensional exponential and logistic maps in chapter 1, the simple extension into the two-dimensional (two-species) case can begin with simple exponential assumptions for each of the populations, namely,

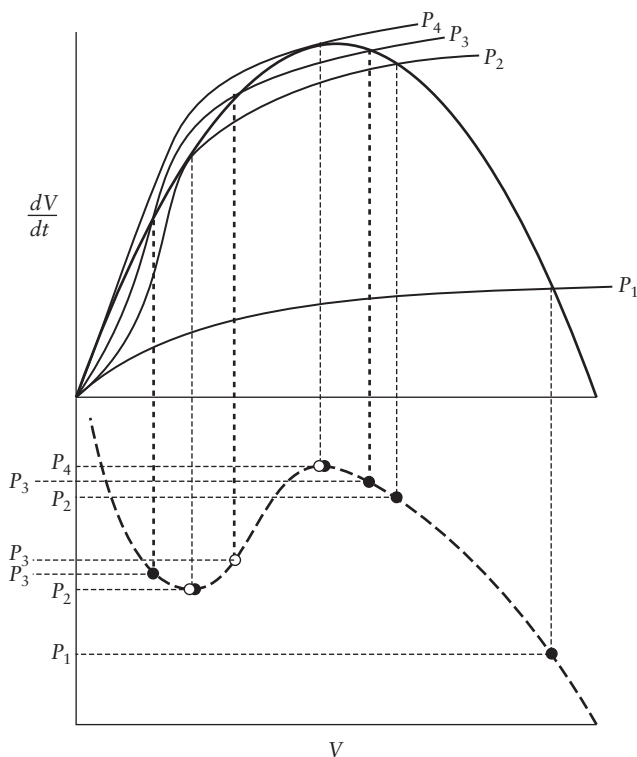


FIGURE 6.17. Constructing the prey isocline from the attractors and repellers of the rate-versus-density graph. The prey rate versus the density graph has been exaggerated so as to illustrate the construction more clearly. The upper graph shows the rate of growth of prey plotted along with four type II functional response curves. Where the rate of growth intersects the functional response is an equilibrium situation, which is then translated to the lower graph for seven particular predator densities (some functional responses result in alternate equilibria). Stable points are indicated with solid circles and unstable points with open circles, and the points are connected with a dashed curve that represents the isocline, with bold dotted lines indicating the points for the one functional response that yields three separate alternative equilibrium points.

$$V_{t+1} = aV_t \quad (8a)$$

and

$$P_{t+1} = cP_t, \quad (8b)$$

where the two parameters (a and c) are the finite rate of increase (see chapter 1).

But two independent exponential equations do not represent a good model of predator-prey interactions because they are not connected. So we follow the procedure we used with the two exponential differential equations representing predator and prey. That is, we first suppose that there is a birth and death process for each of the two populations, namely,

$$V_{t+1} = (b_1 - m_1)V_t \quad (9a)$$

and

$$P_{t+1} = (b_2 - m_2)P_t. \quad (9b)$$

Next we suppose that the death of the prey is directly proportional to the abundance of the predator (i.e., $m_1 = mP_t$), which changes equation 9a to

$$V_{t+1} = (b_1 - mP_t)V_t,$$

and we further suppose that the birth rate of the predator is directly proportional to the abundance of the prey (i.e., $b_2 = bV_t$), which changes equation 9b to

$$P_{t+1} = (bV_t - m_2)P_t.$$

A slight rearrangement of these equations gives

$$V_{t+1} = b_1 V_t - mP_t V_t \quad (10a)$$

and

$$P_{t+1} = bV_t P_t - m_2 P_t, \quad (10b)$$

which is quite clearly a discrete analog to equations 3a and 3b presented earlier in this chapter, the classic Lotka–Volterra equations. Equations 10a and 10b are not the normal way of developing a discrete approach to predator–prey (or host–parasite) interactions, but there are lessons to be learned from this formulation before we go on to the standard treatment.

EXERCISES

- 6.16** Iterate the basic discrete model for 130 iterations with parameters $b_1 = 1.1$, $m = 0.8$, $b = 0.5$, and $m_2 = 0.002$. Plot the time series over time.
- 6.17** An alternative model (to be discussed below) is

$$V_{t+1} = bV_t e^{-aP_t},$$

$$P_{t+1} = cV_t e^{1-aP_t}.$$

Iterate this model for 130 iterations with parameters $c = 1$, $a = 0.3$, and $b = 1.1$. Plot the time series over time. Take logs of the numbers of prey and predators and plot them against one another.

Returning to the basic discrete model (equations 10a and 10b), if we set $V_{t+1} = V_t$ and $P_{t+1} = P_t$, the equilibrium condition is

$$V^* = \frac{1 + m_2}{b}$$

for the predator isocline and

$$P^* = \frac{b_1 - 1}{m}$$

for the prey isocline.

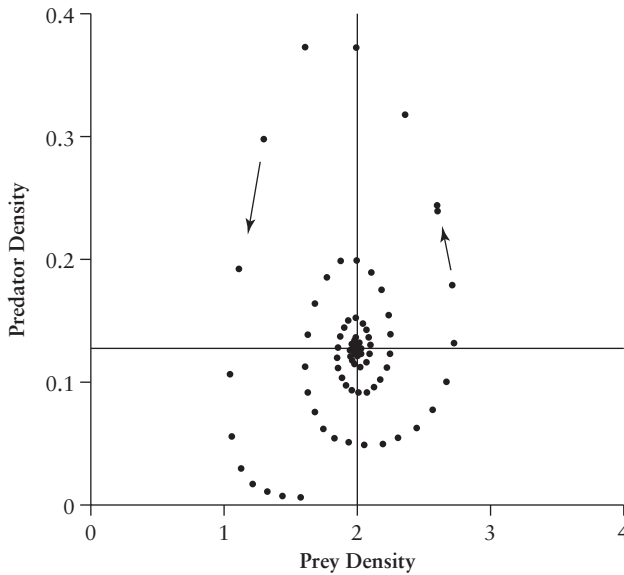


FIGURE 6.18. Results from the discrete form of the Lotka–Volterra predator–prey equations. Note that the equilibrium point is unstable. The horizontal and vertical lines are standard isoclines of the system. The parameter values for this graph are $b = 0.5$, $m = 0.002$, $b_1 = 1.1$, and $m_2 = 0.8$.

The development thus far appears to be quite the same as the continuous case, with the two isoclines simple constant values along each of the axes of the phase plane (the graph of V against P). However, this system does *not* have the “neutral” stability of the original Lotka–Volterra formulation but rather is inherently unstable, although in both cases the qualitative conclusion that the system is fundamentally oscillatory remains true. For example, we show an illustration of the system dynamics in figure 6.18 (compare it to the parallel graph for the continuous time Lotka–Volterra equations in figure 6.2B). Figure 6.18 has points because the formulation is discrete, whereas figure 6.2 has continuous trajectories because the formulation is continuous.

The comparison of figures 6.2 and 6.18 reveals an important fact, not particularly about predator prey interactions but rather about the way in which the theory about them is formulated. The same qualitative formulation that gave us the onerous neutral stability (equations 3a and 3b, figure 6.2B) in continuous time gives us an unstable system (equations 10a and 10b, figure 6.18) in discrete time. Thus the simple fact of converting to a discrete time framework drastically changes a major qualitative characteristic of the system. Indeed a great deal of ecological theory has attempted to deal with this question of the “inherent” instability of the predator–prey (host–parasite) system, and it is well worth remembering that the unstable character of the system stems from use of the discrete rather than continuous platform.

The more classical approach to this problem in discrete time is a bit different. The canonical formulation of this comes from Nicholson and Bailey

(1935), who began by assuming that populations engage in interactions that occur “only during particular periods of time that succeed each other at intervals equal to the length of a generation” (Nicholson and Bailey 1935, 553). Framing the problem in a probabilistic way, let f be the probability that an individual host escapes attack from a predator (a parasitoid), and assume that every host attacked is killed. Thus fV_t is the number of surviving offspring after the “particular period of time” (one generation) ends. Assuming that each of those individuals produces b other individuals (the birth rate), we expect, in the next generation,

$$V_{t+1} = bV_t f. \quad (11a)$$

It is very likely that the probability that a host will escape attack will be a function of the densities of both parasite and host (certainly that of the parasite and perhaps that of the host itself), so to make things as general as possible we let f be a function of both population densities, which is to say, $f = f(V_t, P_t)$.

The parasite will increase its population based on those hosts that did *not* escape (i.e., those that the parasite itself attacks), which is $1 - f$, so that the parasite equation should be

$$P_{t+1} = cV_t(1 - f), \quad (11b)$$

where the parameter c is the conversion rate of the parasite (c number of parasites are produced for each attacked host). Adding the explicit dependence of f on both predator and prey densities, equations 11a and 11b become

$$V_{t+1} = bV_t f(V_t, P_t) \quad (12a)$$

and

$$P_{t+1} = cV_t[1 - f(V_t, P_t)]. \quad (12b)$$

Thus equation 12a simply states that the number of prey next time (V_{t+1}) is equal to the number that escape predation now [$V_t f(V_t, P_t)$] times a reproductive factor (b), and equation 12b states that the number of predators next time (P_{t+1}) is equal to the number of prey that succumb to predation now ($V_t[1 - f(V_t, P_t)]$) times the conversion of prey into predator individuals (c).

A further thought experiment provides a specific form for f . During a single time interval we expect that there will be a certain number of predatory events, V_e . That is, V_e individual prey will be consumed during the time interval. The “crude attack rate” (a), which is defined as the probability that an individual prey will be attacked, can then be expressed as the ratio of the number of prey attacked by an individual predator to the total number of prey present, or

$$a = V_e/V_t,$$

and $V_e = aV_t$ would give us the number of prey attacked if there were only a single predator in the system.

The net attack rate, then, is this number of prey attacked per predator multiplied by the number of predators in the system. Thus we can write

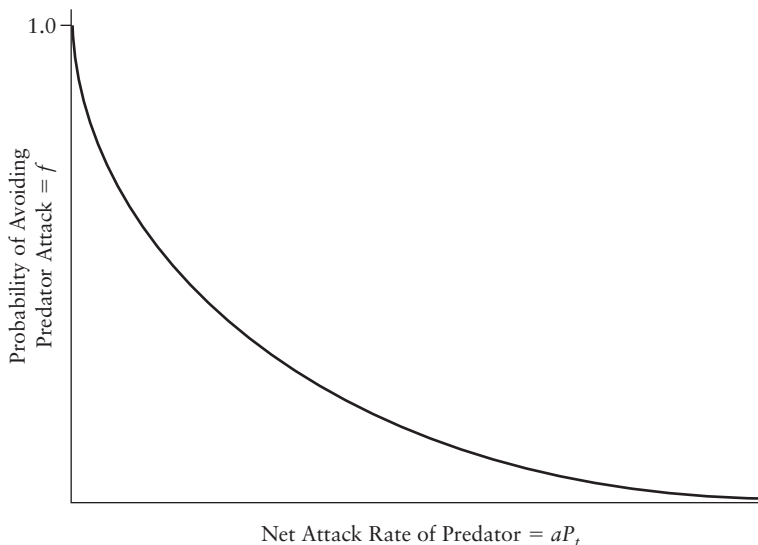


FIGURE 6.19. The probability of avoiding attack by the predator as a function of the net attack rate, aP_t .

$$V_e = aP_t V_t. \quad (13)$$

To avoid confusion, we refer to the parameter a as the crude attack rate (the number of prey consumed per predator) and aP_t as the net attack rate (the total number of prey consumed by the population of predators).

What is likely to be the form of the function $f(V_t, P_t)$, the probability of avoiding predation, as a function of the net attack rate? Certainly it must equal 1.0 when $P_t = 0$ (no predators, no attacks). Furthermore, as P_t approaches a very large number, the likelihood of avoiding attack diminishes toward zero. Thus we expect a relationship something like that presented in figure 6.19. Nicholson's original reasoning was simply to assume that attacks were random and thus that the function f should be proportional to the zero term of the Poisson distribution (the probability of zero attacks; see chapter 5), namely,

$$f = e^{-\left(\frac{V_e}{V_t}\right)}, \quad (14)$$

which is a convenient description of the general form shown in figure 6.19. Equation 13 can be rearranged to give us

$$\frac{V_e}{V_t} = aP_t,$$

and, substituting into equation 14, we have

$$f = e^{-aP_t}. \quad (15)$$

Equation 15 can be substituted into equations 12a and 12b to obtain

$$V_{t+1} = bV_t e^{-aP_t} \quad (16a)$$

and

$$P_{t+1} = cV_t[1 - e^{-aP_t}], \quad (16b)$$

which are the classical Nicholson–Bailey equations (Nicholson 1933; Nicholson and Bailey 1935).

It is clear that these equations are based on a simple and obvious set of assumptions. This fact makes it initially surprising when, much as in the simple elaboration of a discrete Lotka–Volterra approach (as above), the basic equations say that *all* predator–prey (host–parasite) systems will be unstable! In figure 6.20 we present (A) a time series derived from equations 16a and 16b, along with two-phase space diagrams (B and C) illustrating the unstable oscillatory nature of the system.

A great deal of theoretical work has gone into trying to figure out how to stabilize this very simple and popular formulation. Based on previous material in this chapter, it is probably a good guess that adding a density-dependent term to the prey will act in a stabilizing way. To make the simplest modifi-

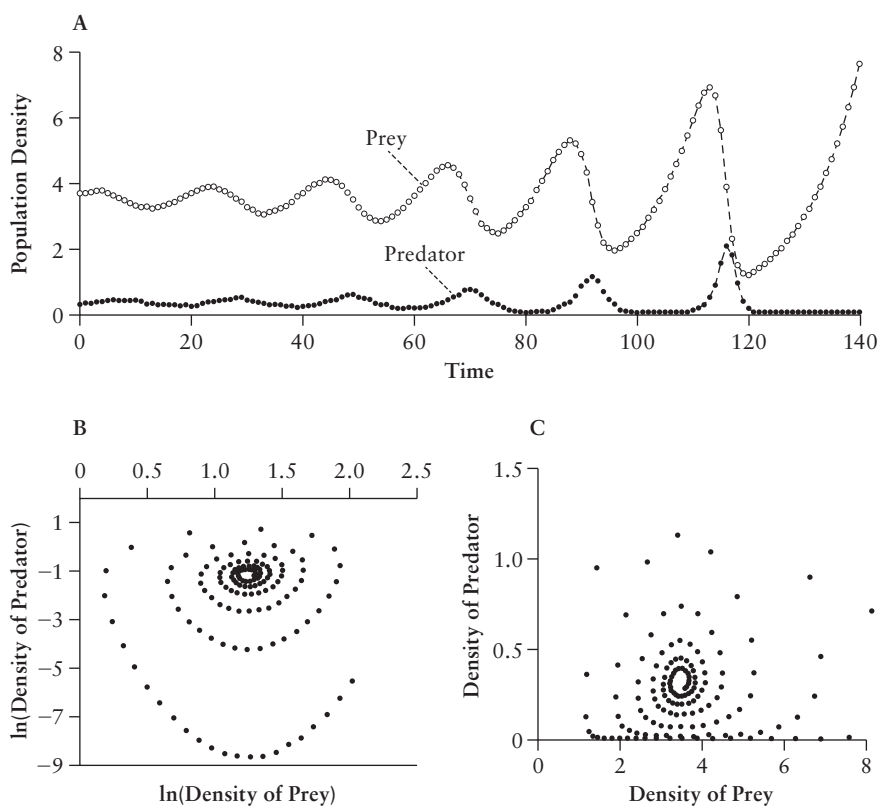


FIGURE 6.20. Solution of equations 16a and 16b with parameters $a = 0.3$, $b = 1.1$, and $c = 1$. At top is a time series (A) derived from equations 16a and 16b. The bottom phase plane plots (B and C) are shown on both the log and the arithmetic scale to emphasize different areas of the graphs.

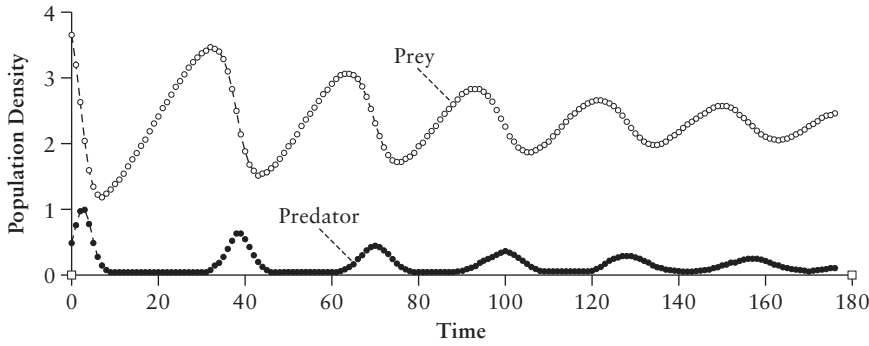


FIGURE 6.21. Solution of equations 17 and 16b with parameters $a = 0.2$, $b = 1.5$, $c = 1.5$, and $K = 40$. The small open circles are the densities of the prey, and the larger closed circles are the densities of the predator.

cation, we simply add a logistic term to the prey equation (equation 16a) to obtain

$$V_{t+1} = bV_t \left(\frac{K - V_t}{K} \right) e^{-aP_t}. \quad (17)$$

Coupling equation 17 with equation 16b and resetting the parameters, the pattern in figure 6.21 is a common outcome, illustrating the switch from unstable oscillations to stable oscillations (i.e., oscillations that damp to a point).

However, it should come as no surprise, given the background information in chapter 4, that adding this sort of nonlinear term to a discrete equation can also lead to a variety of complications. Indeed, with equations 17 and 16b, patterns like that displayed in figure 6.22 also commonly arise. Note that the successive points in time trace out what is effectively a loop (figure 6.22B). Although the position on that loop does not follow the same sort of orderly path that is found in a limit cycle (note the position on the loop of each successive peak of the oscillations in figure 6.22C), all points are constrained to fall exactly on that loop. Such a structure is referred to as an invariant loop.

There is a certain consistency between the continuous formulation and the discrete formulation. However, the distinctions are important. First, as noted earlier, the discrete form is inherently unstable, whereas the continuous form is “neutral” (i.e., neither expanding nor dampening oscillations but rather permanent oscillations that forever return to their starting point—formally, centers), although adding simple density dependence can stabilize it, as also happened in the continuous case (figure 6.4). But this outcome is not universal. With different parameter values, the equivalent of a stable limit cycle can emerge when only density dependence is added, as illustrated in figure 6.23. This does not occur in the continuous case, where density dependence in the prey (alone) always leads to a stable attractor.

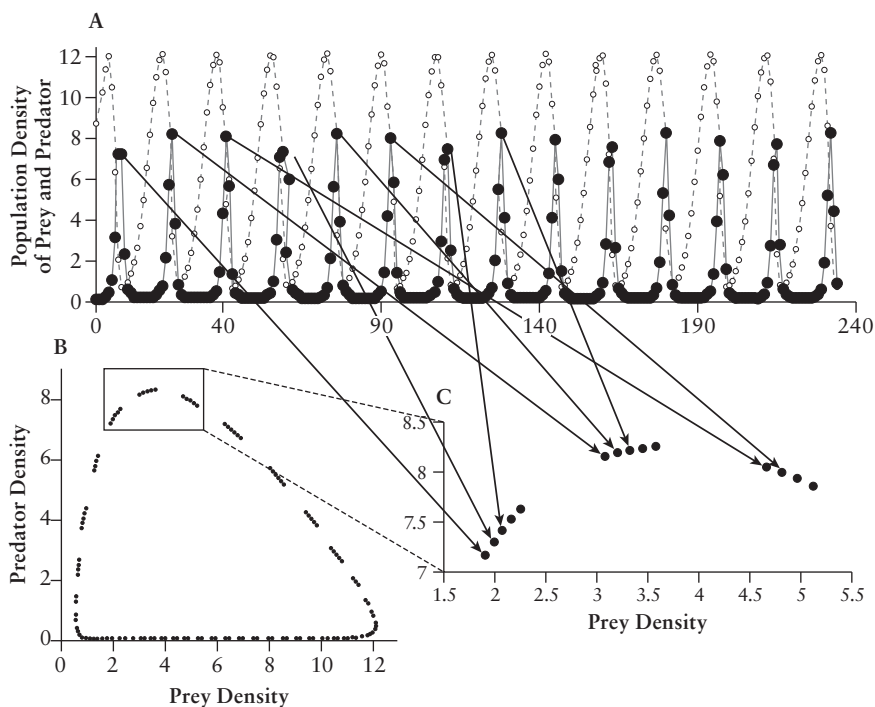


FIGURE 6.22. Solution of equations 17 and 16b with parameters $a = 0.3$, $b = 1.1$, $c = 1.5$, and $K = 50$. (A) Time series illustrating what appears to be qualitatively the same as a stable limit cycle with the predator (solid circles) and prey (open circles) seemingly oscillating on a similar trajectory repeatedly. (B) Phase portrait of the same data as in A, illustrating the smaller groups of points that appear to be forming a loop. (C) Microscopic view of the section in the rectangle of B, showing the particular points that correspond to the points in the time series (A), illustrating the very regular behavior of the points on the loop that is forming. If the model continues for a long period of time, the points form a very dense set that appears to be a closed loop. This structure is referred to as an invariant loop, a commonly observed phenomenon in nonlinear models cast in discrete form.

Although the details are beyond the scope of this text, it is important to note that whereas an invariant loop seems similar to a limit cycle biologically, it is quite distinct mathematically (e.g., note that in figure 6.22C neighboring points do not track to each other). Whether this makes any practical difference depends on the particular application. But in principle it should be noted that the limit cycle-like behavior of an invariant loop emerges from a simple application of density dependence in the prey population and does not require a modification of the predator population, as it does in the continuous formulation of the problem. Furthermore, the combination of nonlinearities and the discrete formulation leads to other complications, as we see next.

Equation 17 is a simple extension that follows the logic originally applied to the exponential map (chapter 1), which is to add a simple quadratic element to the basic exponential idea, leading in chapter 1 to the logistic map.

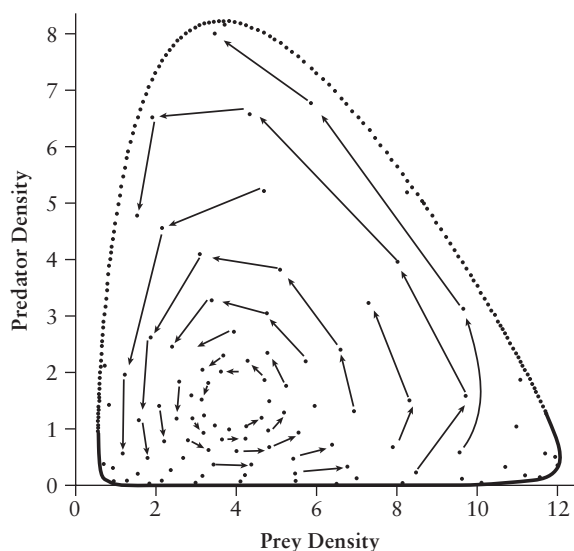


FIGURE 6.23. Dynamic trajectories of the equation set 17 and 16b (with the same parameter values as in figure 6.22). Trajectories initiated near the theoretical equilibrium point are connected by arrows to indicate the dynamical history. Note the unstable oscillations that end at the invariant loop shown as a dashed line (same loop as in figure 6.22).

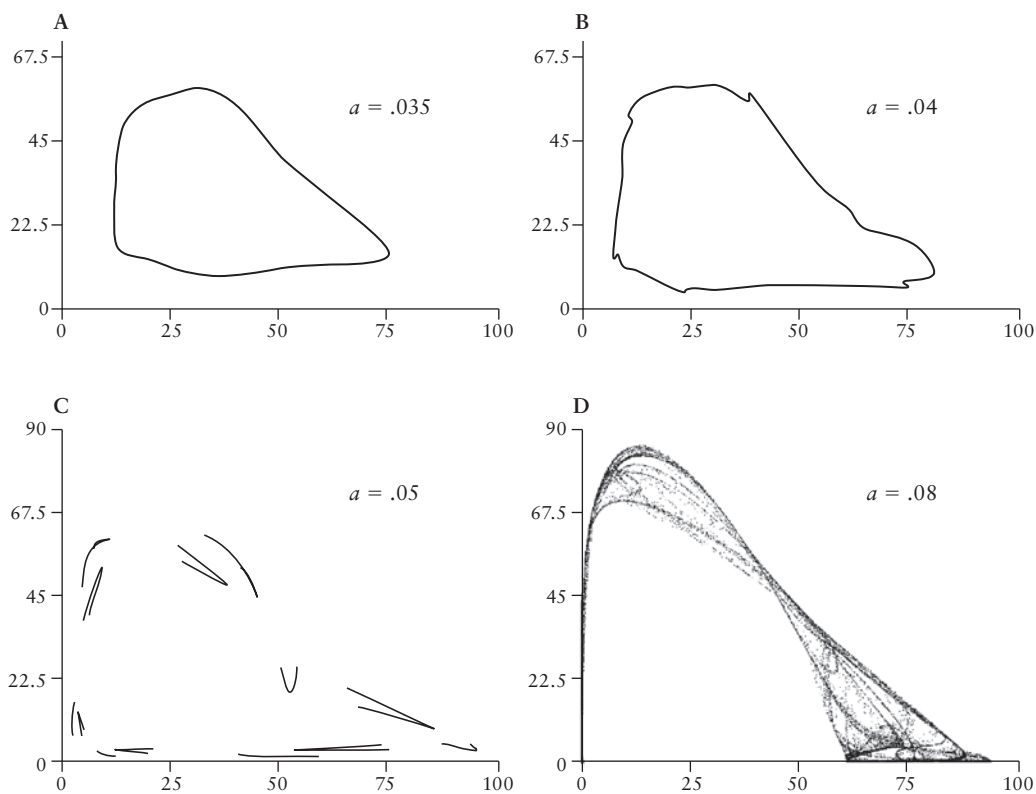


FIGURE 6.24. Phase plane representation of equation set 17 and 16b with changing values of parameter a (the crude predation rate) illustrating (A) an invariant loop, (B) a deformed invariant loop, (C), a fragmented attractor, and (D) a possibly chaotic attractor.

A more popular extension is to incorporate the logic of the Ricker map (see chapter 1) instead, so equation 16a becomes

$$V_{t+1} = bV_t e^{\left(\frac{K-V_t}{K}\right) - aP_t}, \quad (18)$$

in which it is evident that the Ricker equation has just been modified with the addition of $-aP_t$ to the exponent. Formulating the problem with this equation helps to make its relationship with the Ricker equation from chapter 1 clear, but it is not the normal way of presenting it. Rather, if we let

$$b' = be,$$

where e is Euler's constant, equation 18 becomes

$$V_{t+1} = b'V_t e^{\left(\frac{V_t}{K}\right) - aP_t}. \quad (19)$$

Equation set 19,16b is a common representation of the modified Nicholson–Bailey model with prey density dependence added. Its properties are very similar to those of the less well-known system 17,16b.

Using either of these systems, it is possible to demonstrate a rather complicated set of possibilities related not only to the classical instability versus stability issue but also to the generation of invariant loops and their eventual decomposition into chaotic behavior, as illustrated in figure 6.24.

The previous chapter treated the situation in which predator and prey are relatively equivalent in size and life history characteristics, a focus thought to apply to a wide variety of organisms, from parasitic hymenoptera attacking insect hosts to lions attacking zebras. But there is another kind of predation that is so different that it merits an entirely different mathematical approach. When the host is very large, with relatively slow dynamics, and the parasite is extremely small, with very rapid dynamics, the traditional predator–prey approach is not very useful. This is the case of infectious disease, in which the host is usually an animal or a plant and the predator is a microorganism such as a bacterium or a virus (referred to as a pathogen). In this situation we usually assume that the host population is constant and ask questions about the spread of the parasite among host individuals. This is the subject matter of classical epidemiology, and we introduce it here in its simplest form. The interested reader is referred to one of the many excellent texts in epidemiology (e.g. Anderson and May 1991; Bailey 1975; Keeling and Rohani 2008) for more sophisticated development.

EXERCISES

- 7.1 Individuals of the small crustacean *Daphnia pulicaria* float passively in the ocean. A fungal pathogen, *Polycaryum laeve*, regularly attacks them. At any time some individuals in the *Daphnia* population contain the pathogen (are infected), but sometimes they are free of the pathogen (but still susceptible to it). The only way the pathogen can get from one *Daphnia* to another is when the *Daphnia* come in contact with one another. If 25% of the individuals in a particular *Daphnia* population contain the pathogen, what is the probability that a *Daphnia* without the pathogen (an uninfected, or susceptible, individual) will contract them? (Think of the contact process: what are the probabilities of different types of encounters, and what needs to happen for an uninfected *Daphnia* to become infected?)

- 7.2 Repeat exercise 7.1 but with the additional assumption that every time one *Daphnia* encounters another, there is a 50% chance that a pathogen (if present) will move to the uninfected *Daphnia*.
- 7.3 Let the fraction of *Daphnia* that are infected be equal to p and the probability of transferring the pathogen between two *Daphnia* in contact be m . If the rate of change of the fraction of *Daphnia* infected with the pathogen is taken to be equal to the probability of a transfer upon contact, write the general equation for that rate of change.
-

Direct Disease Transmission

We begin by dividing the host population into those infected with the pathogen and those not infected (the susceptibles), in much the same way that we divided the habitats into occupied and unoccupied in the case of metapopulations (see chapter 5). Let I be the number of infecteds and S the number of susceptibles, and suppose that $N = I + S$, where N is the number in the population and remains constant as I and S vary through time. We presume that the pathogen can exist only within a host and that there are no intermediate hosts (this is the definition of a direct transmission system). We assume that the rate of increase of infecteds is proportional to the product of infecteds and susceptibles because the rate of transmission is proportional to the probability that susceptibles will encounter infecteds). This is known as the mass action assumption. Thus we write

$$\frac{dI}{dt} = \beta IS,$$

where the parameter β is the proportionality constant (usually referred to as the transmission coefficient) and essentially represents the probability that if an infected contacts a susceptible the disease will, in fact, be transmitted. Because $N = I + S$, we know that $S = N - I$, so we can substitute for S in the above equation to obtain

$$\frac{dI}{dt} = \beta I(N - I).$$

If we let $N = 1$, I becomes the proportion of the total infected in the population and the model is

$$\frac{dI}{dt} = \beta I(1 - I), \tag{1}$$

which is readily recognizable as the logistic equation, with the carrying capacity equal to unity. Thus, in its simplest manifestation, we expect the time course of a disease to have roughly a logistic pattern, which indeed is frequently observed. For example, Stiven (1967) experimentally introduced a pathogenic microbe (*Hydrumoeba hydroxena*) into a cnidarian (*Chlorohydra viridissima*) and charted the time course of the infection. His results are illustrated in figure 7.1. Note the qualitative pattern, which is identical to the pattern predicted by the simple logistic equation (recall chapter 1).

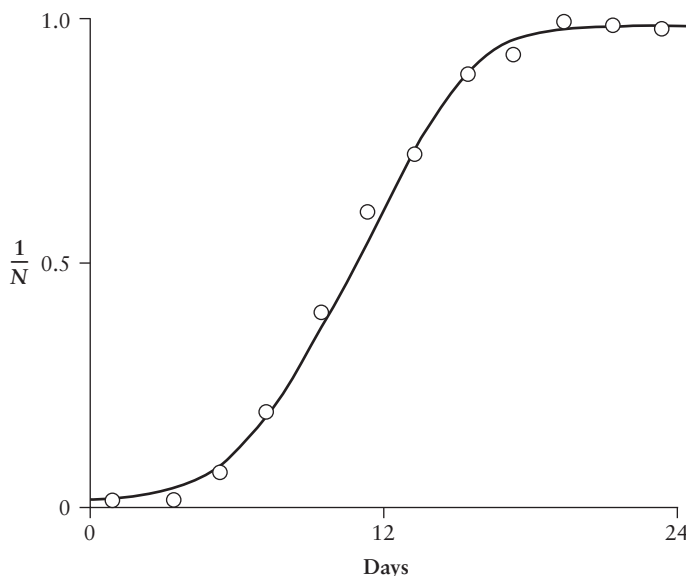


FIGURE 7.1. Time course of an infection. The proportion of a population of the cniderian *Chlorohydra viridissima* infected with the microbe *Hydramoeba hydrox-ena* (from Stiven 1967 as reported in Anderson 1981).

EXERCISES

7.4 The proportions of *Daphnia* infected during a period of 15 days is, in sequence, 0.01, 0.03, 0.03, 0.1, 0.14, 0.3, 0.3, 0.58, 0.7, 0.7, 0.92, 0.92, 0.96, 0.97, 0.99, 0.99. Plot the numbers over time.

7.5 The integrated form of equation 1 is

$$I_t = \frac{I_0}{(1 - I_0)e^{-\beta t} + I_0}.$$

Solve this equation for β , and estimate the value of β for the time series of exercise 7.4.

7.6 Using the value of β you estimated in exercise 7.5, calculate the expected values using the integrated form of equation 1, and plot them along with the time series of exercise 7.4.

7.7 Equation 1 is a perfectly fine representation of the accumulation of infected individuals in the population. What equation would describe the rate of change in the number of susceptibles in the population? What would be the integrated form of that equation (recall exercise 7.5)?

7.8 Begin with a very small infection ($I_0 = 0.001$) and use the integrated form of the mass action equation (the equation in exercise 7.5) to compute the time course for the susceptible population, with $\beta = 1$, then with $\beta = 0.1$. Plot the time series.

7.9 Recall that $I + S = 1.0$, and plot I and S on the same graph for $\beta = 0.9$, then with $\beta = 0.1$.

Not all diseases spread continuously through the host population such that the simple logistic pattern of infection occurs, increasing from almost zero to 100% infected over time. Individual host organisms sometimes are able to shed the disease, that is, become cured of the disease. If the rate of recovery from the disease is ν , we can modify equation 1 to

$$\frac{dI}{dt} = \beta I(1 - I) - \nu I, \quad (2)$$

which you will recognize as equivalent to the basic Levins metapopulation equation from chapter 5. Here β is equivalent to migration (the disease organism “migrates” among hosts) and the ν is equivalent to the extinction rate (the disease organism becomes “extinct” in a particular host). So the equilibrium form is

$$I^* = 1 - \frac{\nu}{\beta}. \quad (3)$$

In disease ecology the ratio ν/β is extremely important. As can be seen from a superficial examination of equation 3, if it is greater than 1.0, the equilibrium proportion of hosts will be negative, which effectively means zero. Thus if the ratio is greater than 1.0, the disease will be unable to spread through the population. On the other hand, if it is less than zero, the disease will become endemic or, if the ratio becomes close to zero, epidemic (i.e., I^* approaches 1.0).

As simple as this model (equations 1 and 2) is and as nicely as some data fit the logistic pattern, most diseases are more complicated and include the concept of resistance or immunity. This is especially true when speaking of human diseases (or, for that matter, the diseases of most vertebrates, some invertebrates, and even some plants, depending on how one defines immunity). Thus, rather than having a population structured with only susceptibles and infecteds, we really have three types, susceptibles, infecteds, and resistant (or recovered and immune), which gives rise to a basic model usually referred to as the SIR model. Originally devised in 1927 by Kermack and McKendrick, the elementary form of the model is captured with a flow chart, as in figure 7.2.

The original idea of Kermack and McKendrick was to study the progress of a disease from the time of initial infection through to its epidemic stage,

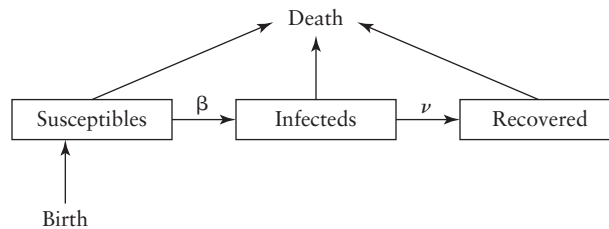


FIGURE 7.2. The basic idea of the SIR model, based on three types of individuals in a population: susceptibles, infecteds, and resistant (or recovered and immune). Note that the process of birth and death are presumed to be perfectly balanced so that the overall population remains constant. Another way of saying the same thing mathematically is that the rate changes between these types are extremely fast compared to the birth and death processes so that we can presume birth and death are balanced.

which, before they developed this model, was not really clear. As the authors originally put it,

As the epidemic spreads, the number of unaffected members of the community becomes reduced. . . . In the course of time the epidemic may come to an end. One of the most important problems in epidemiology is to ascertain whether this termination occurs only when no susceptible individuals are left, or whether the interplay of the various factors of infectivity, recovery and mortality, may result in termination, whilst many susceptible individuals are still present in the unaffected population. (Kermack and McKendrick 1927, 701)

As we shall see, the solution to the problem is quite straightforward with the formation of the simplest model possible.

The equations for susceptibles and infecteds have already been defined by the mass action assumption, namely

$$\frac{dS}{dt} = -\beta SI$$

and

$$\frac{dI}{dt} = \beta IS,$$

in which case it is fairly obvious what the dynamics will be. The infected population grows logistically (as we saw before, replacing S with $1 - I$ in the equation for the infecteds gives the logistic equation). However, adding the compartment of recovered (see figure 7.2) means that we subtract the rate of recovery from the equation for infecteds and add it to the equation for recovered, which results in the following set of equations:

$$\frac{dS}{dt} = -\beta SI, \tag{4a}$$

$$\frac{dI}{dt} = \beta IS - \nu I, \tag{4b}$$

and

$$\frac{dR}{dt} = \nu I. \tag{4c}$$

EXERCISES

7.10 Although it is generally not correct to simply use the discrete approximation of differential equations, sometimes the underlying processes involved in a model can be represented conveniently by doing so. In the case of equation set 4, such is the case (for a limited set of parameter values). For equation set 4 use the discrete approximation of the derivatives (e.g., the first equation would read $S_{t+1} = S_t - \beta S_t I_t$) to compute time series for all three variables (time 0–25) for values of $\beta = 0.8$ and $\nu = 0.3$. (Caution: Because of the discrete approximation, some values may go above 1.0 or below 0. Do

not worry about that; it is a mathematical artifact. Just look at the overall pattern that emerges, and think about the biology.)

7.11 Repeat exercise 7.10 for values of $\beta = 0.6$ and $v = 0.4$.

7.12 Calculate the values of

$$\frac{d \ln(I)}{dt}$$

for values of S ranging from 0 to 1, with $\beta = 0.5$ and $v = 0.3$, and graph them against S . At what value of S will the disease begin to decline?

Equation set 4 is the classic SIR model of epidemiology. Typical solutions for equation set 4 are displayed in figure 7.3.

If we factor I out of equation 4b and do a bit of algebra, we can write

$$\frac{dI}{dt} = vI \left(\frac{\beta}{v} S - 1 \right), \quad (5)$$

and because the largest possible value of S is 1.0, we see that if $\beta/v < 1.0$, the derivative dI/dt will always be negative for any value of I or S , which means that the infected population will always decline. Thus we again see that the ratio β/v takes on special significance. It is usually referred to as the basic reproductive rate and symbolized as R_0 . It must be greater than 1.0 for the disease to become either endemic or epidemic in the host population.

Note that R_0 is the rate of infection per recovery rate, which effectively means the number of new infections produced by a single infection. This is frequently given as the formal definition of the basic reproductive rate, the number of hosts that become infected per single infected host.

If R_0 is greater than 1, we see (refer to equation 5) that the rate of change of I will depend on the value of S . As is evident in figure 7.3B, it is possible

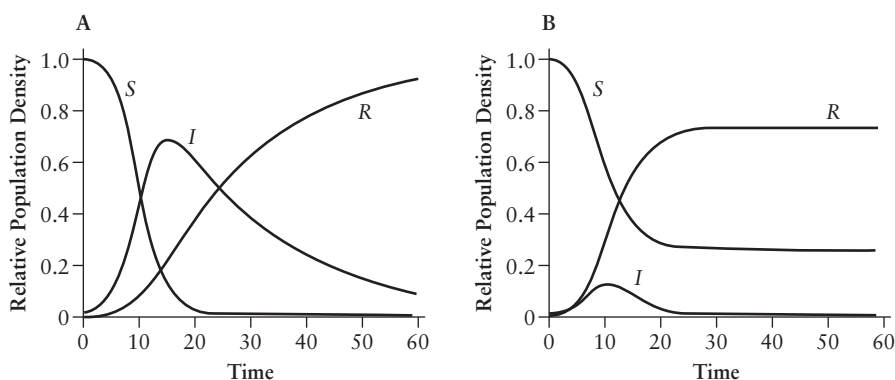


FIGURE 7.3. Typical trajectories from the classical SIR model (equation set 4). (A) With $\beta = 0.5$ and $v = 0.05$, the disease (number of infecteds, I) goes into decline only as the number of susceptibles (S) approaches zero. (B) With $\beta = 0.9$ and $v = 0.5$, note that the infection begins its decline well before the susceptible population approaches zero.

(but not necessarily true) that the disease is effectively eliminated from the population well before all the individuals in the population become infected, because the rate at which new infections occur falls below the rate necessary to sustain the epidemic due to the “shortage” of susceptible individuals. This situation, in which the disease dies out even though not all individuals in the host population have attained immunity, is referred to as herd immunity.

EXERCISES

- 7.13** Use equation 5 or 4b to determine the critical value of S at which the disease will begin to decline. How does that value relate to R_0 ?
- 7.14** At the onset of an infection, public health officials can intervene to attempt to encourage changes in behavior patterns that will decrease host-to-host contact. Mathematically, how will this affect the critical S for which the disease starts declining?
- 7.15** Suppose that individuals can be switched directly from susceptible to resistant without having to become infected. How would that affect equations 4a and 4c?

A convenient graphical vision of the progress of an epidemic can be seen with the use of equation 5. Solving for the critical value of S at the point where I begins to decline, we find that

$$S_{crit} = \frac{v}{\beta} = \frac{1}{R_0},$$

which is shown graphically in figure 7.4.

It is clear that if v/β is very small, the disease is expected to pretty much move through the whole population (indeed, $v = 0$ is the situation described at the beginning of this chapter). That ratio, which is the inverse of the basic reproductive rate, is a useful indicator of how fast the disease will spread

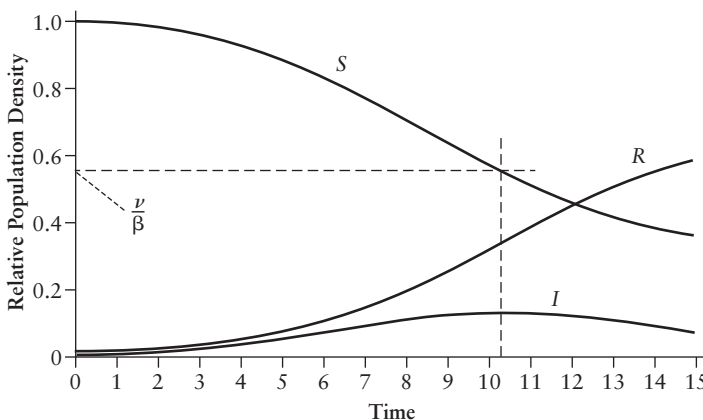


FIGURE 7.4. Expanded view of figure 7.3B illustrating graphically the meaning of the recovery/transmission ratio (i.e., the value of S when I is at its peak).

through the population and how endemic it will eventually become. In figures 7.3B and 7.4, for example, the v/β ratio is about 5.4 (which makes R_0 about 1.85) and results in about 30% of the population never having contracted the disease, a dramatic case of herd immunity. As noted before, if the basic reproductive rate is greater than unity, the disease will persist in the population, whereas if it is less than unity, the disease will disappear. This quantity has been extremely important in understanding the qualitative nature of infections and in some cases in public health planning.

EXERCISES

- 7.16** Some microparasites have increased their transmission rate by evolving the inclusion of another organism to increase the chances of infection. Suppose that a disease organism has an original direct infection rate of β and the chance that two individuals will encounter one another is p . Now suppose that the chance of encounter is increased to $p + \varepsilon$ when an intermediate vector is involved, but the infection rate between vector and human is α . When will it be advantageous for the disease to use the vector?
- 7.17** Suppose that the vector is a mosquito and its rate of biting is b . If the probability that a biting mosquito itself becomes infected is ν , X is the relative number of mosquitoes infected, and Y is the relative number of humans with the disease, write the equation for the accumulation of infected mosquitoes.
-

Indirect Transmission

When another organism is involved in disease transmission, the basic dynamics of infection change at a very fundamental level (Anderson and May 1979). Here we take the case of malaria as a model situation. Other vector-borne diseases are similar but have their own peculiarities, and a model must be constructed for each specific disease. Nevertheless, the underlying structure of the elementary malaria model has historically formed an important point of departure for more complicated and realistic models. Furthermore, malaria affects more people worldwide than any other disease. Ross (1911) was the first to treat this disease from a modeling perspective, and later Macdonald (1957) added considerable realism to the basic Ross model, all of which is summarized by Anderson and May (1991).

We begin with some simplifying assumptions. Suppose that there is no immune response in the host (which is now strongly suspected not to be true), that human victims of the disease do not suffer mortality (which is certainly known not to be true) but rather are always eventually cured of the disease, and finally that the population density of both mosquitoes and humans is constant (i.e., the disease dynamics is very rapid compared to the population dynamics of either host or vector). The first of these assumptions allows us to formulate the problem in terms of two variables, the proportion of the humans who are infected (Y) and the proportion of the mosquitoes that are infected (X). The likelihood of a new human infection is proportional to the probability that the mosquito is infected (X) times the probability that the

human is not $(1 - Y)$, and the rate of loss of old infections through recovery is r . Thus the rate of change of the proportion of host individuals infected can be written

$$\frac{dY}{dt} = aX(1 - Y) - rY, \quad (6a)$$

where a is a proportionality constant. Turning to the dynamics of the vector, the logic is similar in that the rate of acquisition of new infections should be proportional to its biting rate (b), the chance that a mosquito biting an infected host will acquire the disease (v) times the probability of encountering an infected host (Y) times the probability that the biting mosquito is not infected ($1 - X$). (Note that this v is a different parameter from the one with the same symbol used in the classic SIR model; we use it here to be consistent with the large literature on the malaria model.) We must obviously subtract from this the “recovery” rate of the mosquitoes (u) times the density of the mosquitoes (X). Thus we write, for the dynamics of the vector population,

$$\frac{dX}{dt} = bvY(1 - X) - uX. \quad (6b)$$

The parameter a in the equation for human infections can be further developed because it depends on a variety of other parameters. Classically it includes the biting rate of the mosquito (b), the proportion of bites that actually transmit the pathogen (c), the population density of the mosquito (N_Y), and the population density of the humans (N_X). Note that the ratio N_Y/N_X represents the number of mosquitoes per human host, which we can use because it is well known that an individual female mosquito has a constant biting rate, that is, a fixed number of blood meals per unit of time, independent of the population density of the host. So we can postulate that the proportionality term, a , really should be equal to the biting rate times the proportion of bites that actually transmit the pathogen times the number of mosquitoes per human, which gives us

$$\frac{dY}{dt} = bc\frac{N_Y}{N_X}X(1 - Y) - rY \quad (7)$$

as the equation describing the dynamics of host infection. Note that the term bcN_Y/N_X is a single constant (originally the transmission rate a in the previous equation) written in terms of its biologically significant factors. Equations 7 and 6b are basically the same as the original model proposed by Ross (1911).

We analyze the basic model in equations 6a and 6b by looking at the isoclines, i.e., where the derivatives are set equal to zero. The two zero isoclines are

$$X = \frac{rY}{a(1 - Y)}$$

for $dY/dt = 0$ and

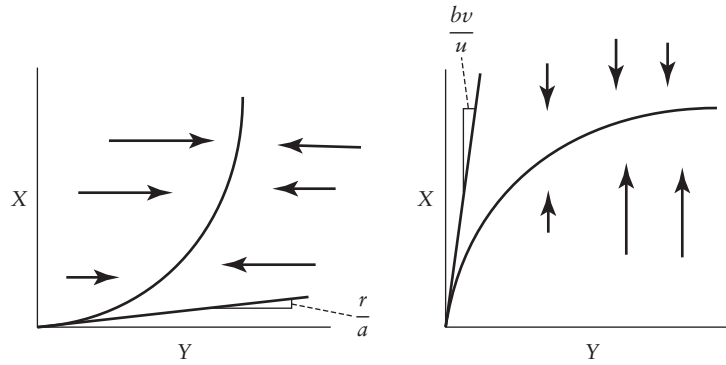


FIGURE 7.5. Isoclines and vector fields for equation set 6. The derivatives (slopes) of each of the isoclines evaluated at the zero point are also illustrated.

$$X = \frac{bvY}{u + bvY}$$

for $dX/dt = 0$. These two isoclines and their vector fields are pictured in figure 7.5.

Putting these two isoclines together, we can graphically analyze the system, as shown in figure 7.6.

This model has two qualitatively distinct outcomes, one in which both populations form a stable equilibrium, the other in which the disease disappears (the number of infected humans and number of infected mosquitoes go to zero). It is of interest to note that when the disease persists (according to this model), the equilibrium state includes less than full infection, which is to say that there will always be some humans and some mosquitoes that are not infected with the disease.

From an examination of the two graphs in figure 7.6, it is clear that their distinguishing feature is the way in which the slopes of the isoclines, evaluated at zero, are related to one another. By inspection of the graphs (and recalling

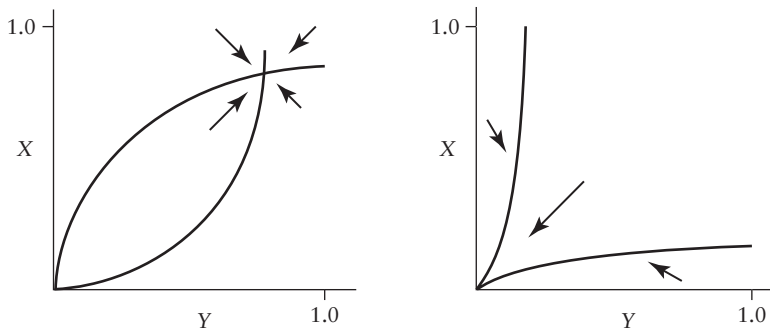


FIGURE 7.6. The two qualitatively distinct arrangements of the isoclines of equation system 6.

the values of those slopes as shown in figure 7.5), we can say that the persistent case will exist when

$$\frac{bv}{u} < \frac{r}{a}.$$

If we let $a = bcN_Y/N_X$ and $m = N_Y/N_X$ (from equation 7) and rearrange a bit, we see that this condition becomes

$$\frac{mb^2cv}{ru} < 1.0.$$

In the classical malaria literature the left-hand side of this equation is usually symbolized as z_0 and is the basic reproductive rate of the disease. A bit of reflection will convince the reader that this measure is analogous to the basic reproductive value defined earlier for a directly transmitted disease. In the history of public health dealing with malaria, the basic reproductive rate has been an important measure. Strategies need to be devised to push z_0 as low as possible.

8 Competition

In the nineteenth century, the notion of competition between different species of animals and plants was common. Indeed, competition was part of the intellectual context in which Charles Darwin and Alfred Russel Wallace formulated the theory of evolution through natural selection. Most frequently the idea of competition was thought of more or less as a sports metaphor. Two teams compete, one wins. The idea seems to have been too obvious to actually write about or think through too clearly. But that all changed when Lotka (1926), Volterra (1926), and Gause (1934) expanded on the elementary ideas of density dependence.

EXERCISES

8.1 The basic logistic equation,

$$\frac{dN}{dt} = rN \frac{K - N}{K},$$

says that the per capita rate of growth of a population will decrease as a linear function of the density of that population. The biological interpretation of

$$\frac{K - N}{K}$$

is the fraction of the potential environment still available to the population. Suppose that $K = 100$ and $r = 1.2$. Make a graph of dN/dt as a function of N for the logistic equation. Now suppose that another species comes along and takes up a quarter of the space originally available. How will the graph change? Repeat for a species that takes up half of the original space.

8.2 Repeat exercise 8.1, but assume that there is a second species that is variable, so that X individuals of this second species will take up X parts of the original space. If you

fix N at 50, what will be the relationship between dN/dt and X ? First make a graph of dN/dt , and then write the equation describing the situation.

- 8.3** Suppose that each individual of the other species (call its density X) takes up only half of the space taken up by each individual of the first species (call its density N). Repeat exercise 8.2 but with this new information, and make the graph of dN/dt and compare it with the graph of exercise 8.2.
- 8.4** For the information in exercise 8.3, write the differential equation describing dN/dt as a function of N and X . Set the derivative equal to zero (finding the values for which N will not change), and graph N versus X . Suppose the second species (X) takes up 90% of the environment of the first species (N). Draw the graph, and compare it with the graph of the first case.
-

Competition: First Principles

At its most elementary level, interspecific competition involves two species using a similar resource. It rapidly gets more complicated, but stripping the phenomenon of all its complications, this is the basic principle—two consumers consuming the same resource. If we suppose that the resource is a self-replicating, dynamic one (we treat other cases later), we can treat it as a logistic population with two consumers using the resource at two different rates. Thus the resource equation is

$$\frac{dR}{dt} = rR \left[\left(\frac{k - R}{k} \right) - a_1 N_1 - a_2 N_2 \right], \quad (1)$$

where R is the density (or biomass) of the resource, r is its intrinsic rate of increase, k is the carrying capacity of the resource, a_i is the rate at which the resource is consumed by the i th consumer species, and N_i is the population density of the i th consumer species. Now let's suppose that each of the consumer populations behaves according to the simple Lotka–Volterra predator–prey equations (see chapter 6), such that we have

$$\frac{dN_1}{dt} = (b_1 R - m_1) N_1 \quad (2a)$$

and

$$\frac{dN_2}{dt} = (b_2 R - m_2) N_2, \quad (2b)$$

where the birth rate is proportional to the resource density, such that b_i is the predation rate (consumption rate) and m_i is the death rate of the i th consumer species. We now presume that the dynamics of the resource is very rapid compared to that of the consumer, such that we can set the derivative of equation 1 equal to zero and solve for the equilibrium value of R . After some algebraic manipulation we obtain

$$R^* = k(1 - a_1N_1 - a_2N_2).$$

Because we have assumed that the resource dynamics are very fast compared to the consumer dynamics, we can approximate the consumer dynamics at any particular value of N_1 and N_2 by simply substituting R^* in place of R in equations 2a and 2b. Thus we have

$$\frac{dN_1}{dt} = [b_1k(1 - a_1N_1 - a_2N_2) - m_1]N_1 \quad (3a)$$

and

$$\frac{dN_2}{dt} = [b_2k(1 - a_1N_1 - a_2N_2) - m_2]N_2, \quad (3b)$$

and by rearranging we can write

$$\frac{dN_1}{dt} = (b_1k - m_1) \left[\frac{\frac{b_1k - m_1}{b_1a_1k} - N_1 - \frac{a_2}{a_1}N_2}{\frac{b_1k - m_1}{b_1a_1k}} \right]$$

and

$$\frac{dN_2}{dt} = (b_2k - m_2) \left[\frac{\frac{b_2k - m_2}{b_2a_2k} - N_2 - \frac{a_1}{a_2}N_1}{\frac{b_2k - m_2}{b_2a_2k}} \right].$$

If we make the substitutions

$$b_ik - m_i = r_i,$$

$$\frac{b_ik - m_i}{b_ia_ik} = K_i,$$

and

$$\frac{a_j}{a_i} = \alpha_{ij},$$

we obtain

$$\frac{dN_1}{dt} = r_1N_1 \left[\frac{K_1 - N_1 - \alpha_{12}N_2}{K_1} \right] \quad (4a)$$

and

$$\frac{dN_2}{dt} = r_2N_2 \left[\frac{K_2 - N_2 - \alpha_{21}N_1}{K_2} \right], \quad (4b)$$

which are the classic Lotka–Volterra competition equations. K_i refers to the carrying capacity of the i th consumer species, while k refers to the carrying capacity of the resource. Note the term in brackets. It is very similar to—

indeed, it is conceptually the same as—the density-dependent term in the logistic equation, except that another species is involved. The α_{ij} s are referred to as the competition coefficients, and it can be seen here that they represent the use of resources by the competing species (j) relative to the use of resources by the species whose dynamics are being considered (i). Note that they do *not* represent the absolute intensity of competition but are rather a ratio of the intensity of interspecific to intraspecific competition.

The equations above dealt with a single resource. Similar developments with two distinct resources become algebraically more cumbersome and are presented in a later section. These alternative developments are important in the sense that interspecific competition is almost never for a single resource (indeed, as described below, most ecologists think that two species competing for a single resource cannot persist together forever). So the definition of the competition coefficient is not really as restrictive as it appears above.

EXERCISES

- 8.5** Find the equilibria for equations 4a and 4b (set the derivatives to zero and solve), and represent both in terms of N_1 as a function of N_2 .
- 8.6** Plot both the equations you derived in exercise 8.5 for $\alpha_{12} = 0.8$, $\alpha_{21} = 0.4$, and $K_1 = K_2 = 100$ on the same graph. If you draw a line from the carrying capacity of the first species to the carrying capacity of the second species, what would that mean with regard to the point at which the two isoclines would cross?
- 8.7** Plot both of the equations you derived in 8.5 for $\alpha_{12} = 0.8$, $\alpha_{21} = 0.6$, and $K_1 = K_2 = 150$ on the same graph. Now redraw the equation $N_1 = K_1 - \alpha_{12}N_2$ on the same graph for $K_1 = 200$, and repeat for $K_1 = 100$ and $K_1 = 50$.
- 8.8** Repeat exercise 8.7 for $\alpha_{12} = 1.3$ and $\alpha_{21} = 1.2$.
-

An alternative derivation of the classic equations follows a more phenomenological path based on the simple logistic equation. Namely, begin with the two logistic equations,

$$\frac{dN_1}{dt} = r_1 N_1 \left(\frac{K_1 - N_1}{K_1} \right) \quad (5a)$$

and

$$\frac{dN_2}{dt} = r_2 N_2 \left(\frac{K_2 - N_2}{K_2} \right), \quad (5b)$$

representing the logistic growth of two populations in isolation, where N_i is the population density of the i th species, r_i is the intrinsic rate of natural increase of the i th species, and K_i is the carrying capacity of the i th species. Recall the biological interpretation of the term $(K - N)/K$ as the proportion of environmental space (or critical resources or places to hide, etc.). When another species is competing for that environmental space, the term must

include that other species. Thus, rather than $(K - N)/K$, it is customary to write $(K_1 - N_1 - \alpha_{12}N_2)/K_1$ for species 1 and a parallel term for species 2. This term simply expands the general idea presented in the logistic equation, conceptualizing the second species as not quite equivalent to the first, which is what the α term, the competition coefficient, is all about. This coefficient effectively converts an individual of species 2 to the equivalent of an individual of species 1 and can be thought of as the ratio of interspecific to intra-specific competition. For example, suppose that the environment of species 1 includes 100 places to hide. Then $K_1 = 100$. Suppose that a single individual of species 2 needs one and one-half of these places to hide. Then the value of α_{12} will be 1.5, indicating that two individuals of species 2 have the same effect on the environment (and thus on available resources) as three individuals of species 1 (i.e., $3/2 = 1.5$). This stylized argument is presented pictorially in figure 8.1.

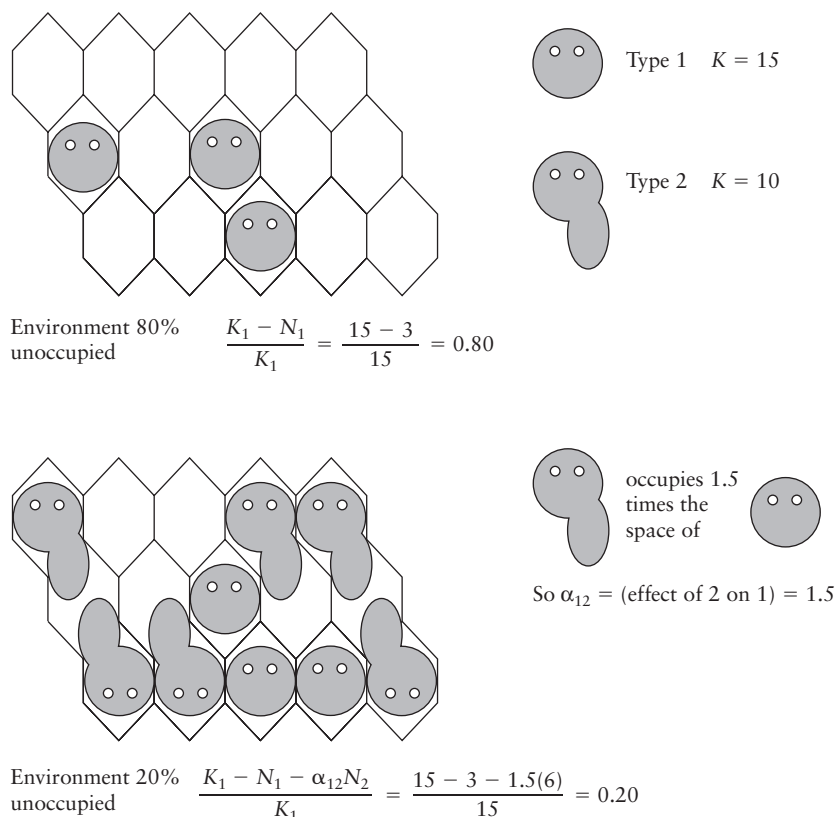


FIGURE 8.1. Diagrammatic representation of the meaning of carrying capacity and competition coefficient. The environment for organism type 1 contains 15 places. Organism type 2 occupies one and one-half of the environmental places that type 1 could occupy. So K for type 1 is 15, and K for type 2 is 10. Because type 2 occupies 1.5 of the environmental space potentially occupied by type 1, the competition coefficient is 1.5.

Equations 1a and 1b are thus modified to

$$\frac{dN_1}{dt} = r_1 N_1 \left[\frac{K_1 - N_1 - \alpha_{12} N_2}{K_1} \right] \quad (6a)$$

and

$$\frac{dN_2}{dt} = r_2 N_2 \left[\frac{K_2 - N_2 - \alpha_{21} N_1}{K_2} \right], \quad (6b)$$

which are the classic Lotka–Volterra competition equations.

Isocline Analysis of the Lotka–Volterra Competition Equations

The traditional way of analyzing these equations is by examining the “isoclines” and deducing the qualitative behavior of the model (as we have already done with predator–prey models). The isocline is the set of all points for which the derivative is exactly zero, which is to say the set of points for which the tendency of the population to increase is exactly balanced by the tendency to decrease. Thus, setting the derivative of one of the equations equal to zero gives the equation for all the combinations of density of species 1 and species 2 for which the state variable in question (N_1 or N_2) does not change. If we set the derivative of equation 6a equal to zero, we obtain

$$0 = r_1 N_1 (K_1 - N_1 - \alpha_{12} N_2) / K_1.$$

Dividing both sides of the equation by $r_1 N_1 / K_1$, we obtain

$$0 = K_1 - N_1 - \alpha_{12} N_2,$$

which can be rearranged as

$$N_1 = K_1 - \alpha_{12} N_2, \quad (7)$$

which is a linear equation in the variables N_1 and N_2 . If the above operations had been done with an inequality rather than an equality, we would easily see that if

$$\frac{dN_1}{dt} > 0,$$

then

$$N_1 < K_1 - \alpha_{12} N_2. \quad (8)$$

Thus we have the important result that if relation 8 is true, the first population must be increasing. Using the same reasoning, we see that if

$$\frac{dN_1}{dt} < 0,$$

then

$$N_1 > K_1 - \alpha_{12} N_2. \quad (9)$$

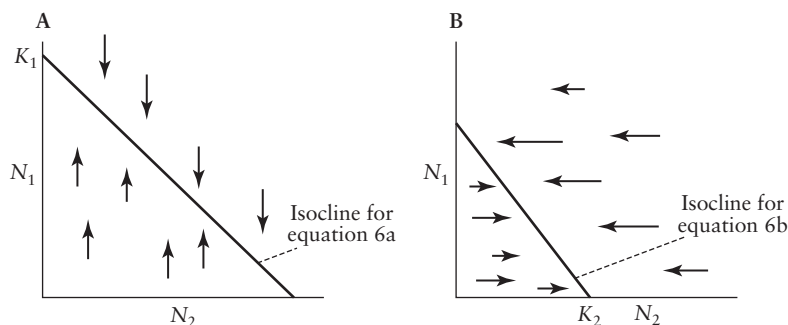


FIGURE 8.2. Isoclines of the competition equations (vectors indicate qualitative behavior on either side of the isoclines).

If relation 9 is true, the first population must be decreasing. Examining relations 7, 8, and 9 on a graph (figure 8.2A), we immediately get an idea of the dynamic behavior (that is, the changes through time) of the first population. Any population above the isocline will decline (relation 9), whereas any population below the isocline will increase (relation 8). Precisely the same analysis can be done for the second species, as presented in figure 8.2B. Figure 8.2 plots the density of the two competing species (the two state variables) against each other and is referred to as the phase plane.

Our main interest in analyzing competition is understanding how the two populations interact when they are together. So it is necessary to put the two together by examining the dynamics represented in figures 8.2A and B on the same phase plane. In figure 8.3A the two graphs from figure 8.2 are superimposed on one another, illustrating the vectors (arrows) moving parallel with either the abscissa or the ordinate, indicating a change in either one or the other species. But to examine the simultaneous changes in both species, we must sum these two vectors, that is, determine the direction in which the two populations will go when simultaneously changing. This is done in figure 8.3B, which illustrates how, qualitatively, the vectors are summed.

The diagram in figure 8.3B is a convenient tool for viewing the general dynamics of the system. It is relatively easy to imagine two populations beginning at an arbitrary point on the graph and being driven by the underlying biological rules in the directions indicated by the vectors. In this particular case we see that the ultimate outcome of competition is the extinction of species 2 and the total dominance of species 1. Analyzing the dynamics from a global perspective (that is, looking at the entire graph, not just the final equilibrium point), it is clear that the system has two repellers (one at K_2 and one at 0), and one attractor (at K_1).

The classic analysis of competition has four general qualitative outcomes based on the fact that you can place linear isoclines on the phase plane in four qualitatively distinct patterns, as shown in figure 8.4. Figure 8.4A is the

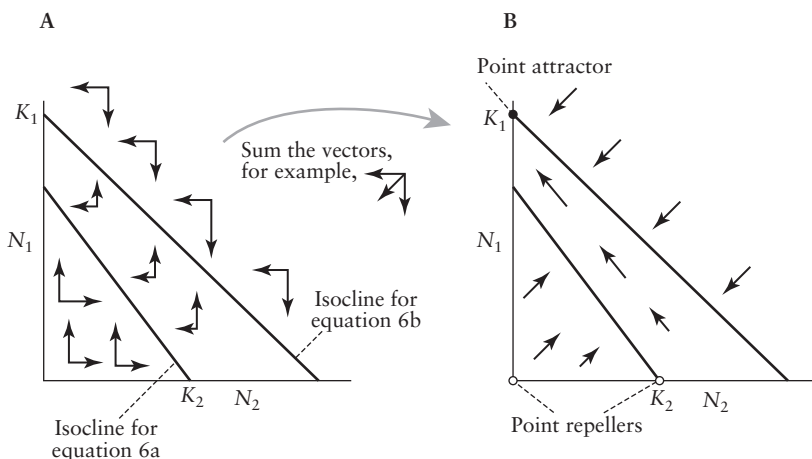


FIGURE 8.3. (A) Putting both isoclines on the same graph and (B) summing the vectors to get the vector field of the two-dimensional system.

same as figure 8.3B and has already been discussed. Figure 8.4B is the reverse of the situation in 8.4A, in which species 2 is the winner of the competition. Thus figures 8.4A and B represent what some would intuitively think of as the typical competition process, in which either species 1 is the better competitor and wins or species 2 is the better competitor and wins. The sports metaphor applies well.

A slightly different situation is presented if the parameters are changed so that the two species have strong effects on the other species relative to effects on themselves (both the competition coefficients are large), as shown in figure 8.4C. Here there will be a winner and a loser, but which species wins and which loses depends on where the ecosystem begins. Whichever species has the advantage at the beginning will generally win in the end. Where the two isoclines cross is an equilibrium point by definition, because it is where the rate of change of both species is exactly equal to zero. But an examination of the vector field (see figure 8.4C) shows that vectors move away from that point, meaning that it is a repeller. It is a slightly different form of repeller than we have seen before, because some vectors point toward it and other vectors point away from it. Rather than a simple hill or a valley (refer to the physical models in chapter 4), the appropriate physical model is a saddle, as we already discussed in chapter 4 and as is illustrated here in figure 8.5. A marble balanced exactly in the middle of the saddle will stay there unless perturbed, in which case it will fall off the side of the saddle. It is also worth noting that this arrangement is highly sensitive to initial conditions (just as is a chaotic system) in that a marble released from two apparently similar points may wind up at totally different attractors, which is to say that very small changes in the point of initiation can have large effects on the future trajectory of change.

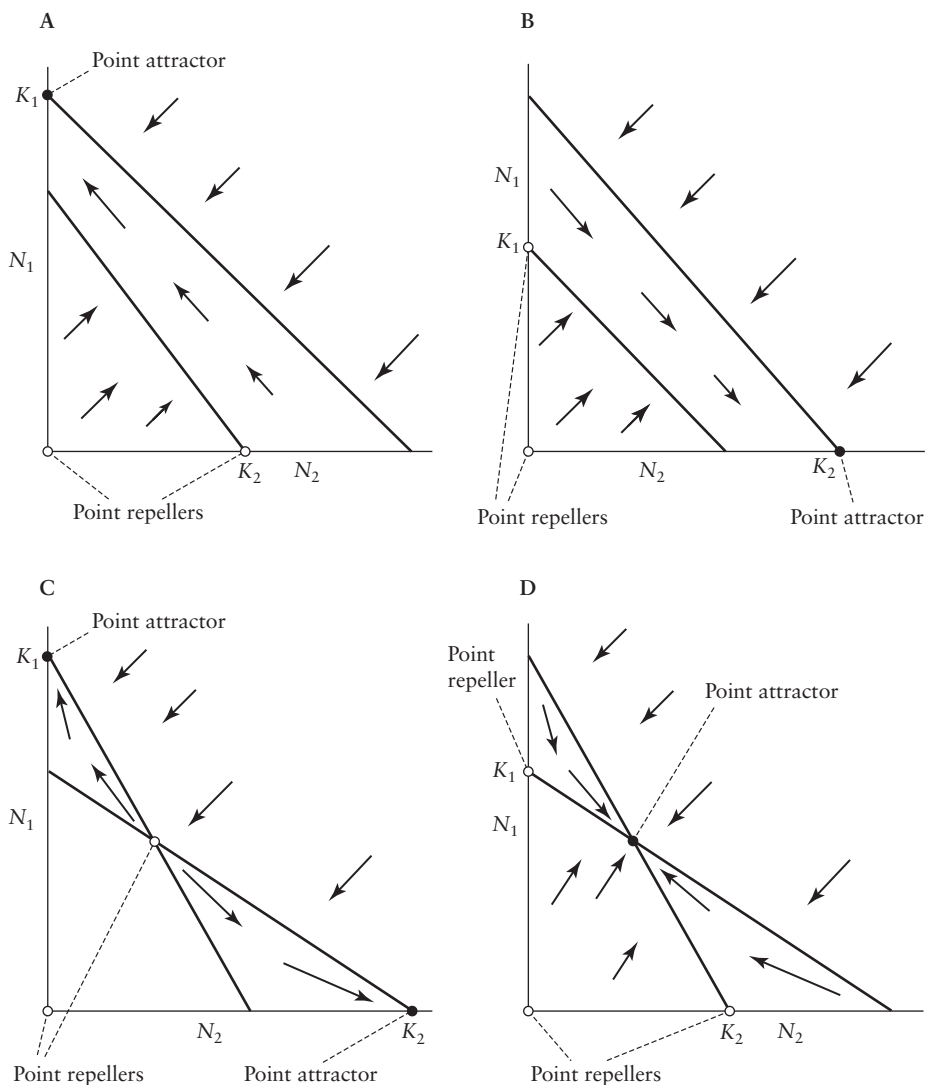


FIGURE 8.4. The four classic cases of interspecific competition. (For meanings of panels, see text.)

A final situation, perhaps the most important, is portrayed in figure 8.4D. An examination of the vector field shows that the equilibrium point at which the two isoclines cross is an attractor. This is also the only case in which the attractor includes both species at positive densities, that is, the only attractor that does not involve the local extinction of one or the other species. This is the case of competitive coexistence.

An important distinction needs to be made between the graphs in figure 8.4A, B, and C (where one of the species is driven to local extinction) and the graph in figure 8.4D (where both species exist together in perpetuity). In

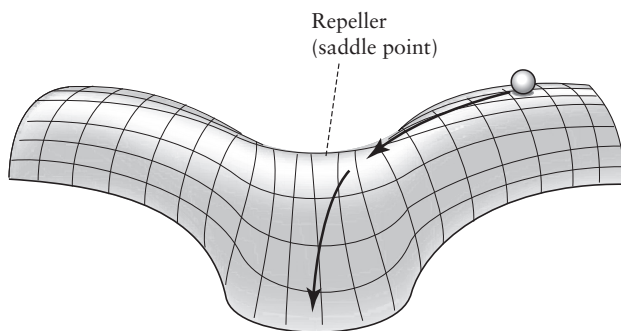


FIGURE 8.5. Physical model illustrating the behavior of a saddle point repeller.

the first three cases, the outcome is similar to that of the metaphorical sports match, where we expect a winner and a loser. But the last case is effectively a tie. A better metaphor is a chess match in which legitimate outcomes are either checkmate or stalemate, the first occurring when there is a clear winner, the last when the players have played to the end and are tied. The difference between the two cases forms the basis for one of the most important principles in ecology, the competitive exclusion principle.

EXERCISES

- 8.9** Suppose that there are two competing species with $K_1 = 95$, $K_2 = 80$, $\alpha_{12} = 1.1$, and $\alpha_{21} = 0.75$. Graph the isoclines, and identify (graphically) the stable and unstable equilibrium points. Repeat for $K_1 = 95$, $K_2 = 80$, $\alpha_{12} = 0.75$, and $\alpha_{21} = 1.1$ and for $K_1 = 100$, $K_2 = 95$, $\alpha_{12} = 1.2$, and $\alpha_{21} = 1.1$.
- 8.10** Repeat exercise 8.9 for $K_1 = 95$, $K_2 = 80$, $\alpha_{12} = 1.1$, and $\alpha_{21} = 0.75$ and for $K_1 = 100$, $K_2 = 95$, $\alpha_{12} = 1.2$, and $\alpha_{21} = 1.1$. On each graph sketch in the equation $N_2 = K_2 - (K_2/K_1)N_1$. Where does the equilibrium point lie in each case?

The competitive exclusion principle in its most elementary form has quite a simple graphical interpretation. In figure 8.6 we illustrate the two most important cases, the case of competitive exclusion in figure 8.6A and that of competitive coexistence in figure 8.6B. The equilibrium point occurs where the isoclines cross, as always. If a line is drawn between the two carrying capacities (the dashed line between K_1 and K_2 in figure 8.6), the position of the equilibrium point either above or below that line differentiates between coexistence and exclusion. In figure 8.6A the point falls below the line connecting the two carrying capacities, and exclusion will take place. In figure 8.6B the point falls above the line connecting the two carrying capacities, and coexistence will take place.

The underlying reason for exclusion or coexistence can be readily seen in these graphs. The reduction of density from carrying capacity to the equilibrium point is illustrated on the two axes for figure 8.6A and B. In the case of

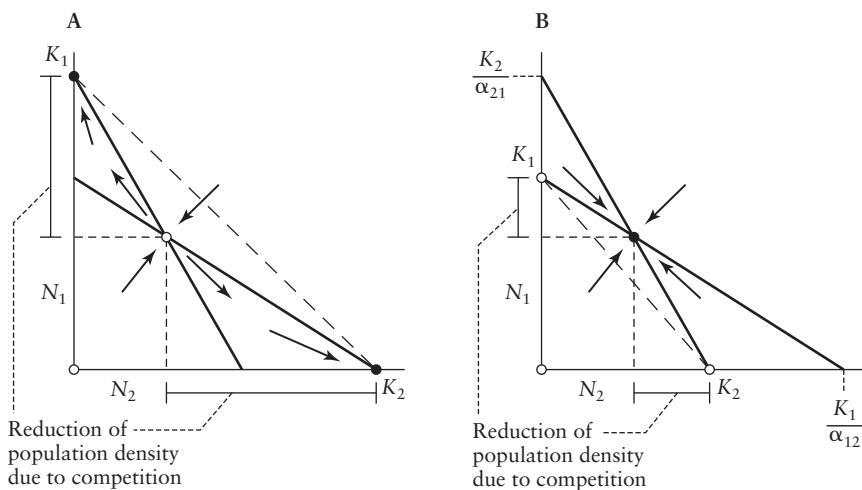


FIGURE 8.6. Explanation of the competitive exclusion principle. (A) Competitive exclusion. (B) Competitive coexistence.

exclusion (figure 8.6A), the reduction has been relatively large, whereas in the case of coexistence, the reduction has been relatively small. So we can qualitatively state the principle that if the effects of interspecific competition on equilibrium density are too great, exclusion will occur, and, most important, if the effects of interspecific competition on equilibrium density are relatively weak, both species can live together in perpetuity, despite the fact that they are competing.

EXERCISES

- 8.11** On a coffee farm in southern Mexico, ants were sampled at five different heights in the vegetation, 0, 1, 2, 3, and 4 meters (0 was on the ground). Three species of ants were encountered, *Pheidole protensa* (P), *Pheidole synanthropica* (S), and *Crematogaster carinata* (C). The proportional distribution of all the sampling sites at which a particular species was encountered was as follows:

Height	P	S	C
0	90	30	0
1	10	55	0
2	0	13	2
3	0	2	18
4	0	0	80

Graph the abundances as a function of position in the vegetation for all three species. What can you hypothesize about the competitive interaction among the three species based on these observations?

- 8.12** Two other sites were located on the same farm, one at which C was absent and one at which S was absent. The same classes of data for these two sites were as follows:

Height	<i>P</i>	<i>S</i>	<i>C</i>
0	90	5	0
1	10	10	0
2	0	15	0
3	0	30	0
4	0	40	0

and

Height	<i>P</i>	<i>S</i>	<i>C</i>
0	70	0	0
1	25	0	0
2	5	0	2
3	0	0	18
4	0	0	80

Graph the abundances as a function of position in the vegetation for all three species at both sites. What can you now hypothesize about the competitive interactions among the three species using the above data and the observations from exercise 8.9?

Niches and Competitive Coexistence and Exclusion

The competitive exclusion principle is frequently related to the idea of the ecological niche, which itself has a long history in ecological thought. Recall that the competition coefficient, α_{ij} , is defined as the conversion of individuals of species *j* into individuals of species *i* (the species whose dynamics is being considered) and therefore represents the intensity of interspecific competition relative to that of intraspecific competition. As such, it combines two phenomena: the degree of overlap in niches (i.e., the degree of overlap in requirements for survival and reproduction) and the species' ability to compete for the shared resources (which is related to its consumption rate). If the species have equal competitive abilities, if the degree of overlap is high, it is likely that each will have similar effects on the other species as on themselves and so have competition coefficients close to one; similar values of α_{ij} slow down the dynamics of competitive exclusion, but one species (the one with the larger α_{ij}) will still win at equilibrium in this model. If the species' niches are distinct, it is likely that they will have smaller competition coefficients because of less overlap in resource requirements (a lower level of interspecific competition relative to intraspecific competition), and coexistence will therefore occur. Fish in a lake do not compete with mice living around the lake even though their ranges, as plotted on a map, overlap. Their niches are obviously different. But fish may compete with tadpoles because both may eat the algae in the lake, although the tadpoles may eat smaller items than the fish, and therefore their niches are not exactly alike. But if two species of mice eat exactly the same food, live in exactly the same habitat, have exactly the same nesting requirements, and so forth, they are likely to compete intensely because their niches are so similar. One or the other species would likely be

excluded from the region at equilibrium. Moving from the fish–mouse competition (with very low niche overlap) to the fish–tadpole competition (with intermediate niche overlap) to the mouse–mouse competition (with complete niche overlap), the competition coefficients are likely to go from zero to some high value, and the likelihood of competitive exclusion at equilibrium will likewise go from zero to some high value. To the extent that the niches are different, the likelihood that exclusion will occur goes down, and to the extent that the niches are the same, that likelihood will go up. This phenomenon is frequently loosely summarized by noting that no two species can occupy the same niche and coexist (this point is treated more completely in a later section of this chapter).

More formally, note that the intercepts in figure 8.6 must be related to one another in a specific way to differentiate between exclusion (figure 8.6A) and coexistence (figure 8.6B). As indicated in figure 8.6B, the intercepts of the first isocline (the isocline for the first species) are K_1 (on the ordinate) and K_1/α_{12} (on the abscissa). The intercepts of the second isocline are similar with appropriate adjustment of subscripts. (You can derive these intercepts by manipulating the equations for the isoclines.) The two graphs in figure 8.6 suggest that a specific arrangement of the intercepts is required for coexistence (figure 8.6B). In particular, we require that the intercept of the first isocline on the abscissa be greater than the intercept of the second isocline on the abscissa and that the intercept of the second isocline on the ordinate be greater than the intercept of the first isocline on the ordinate. That is,

$$K_1/\alpha_{12} > K_2$$

and

$$K_2/\alpha_{21} > K_1$$

are the two conditions that must be satisfied if the two species are to coexist. With a small amount of algebraic manipulation we see that those conditions can be written

$$\alpha_{12} < K_1/K_2 \tag{10a}$$

and

$$\alpha_{21} < K_2/K_1. \tag{10b}$$

A convenient interpretation is seen if the species have the same carrying capacities: both species must have competition coefficients less than 1; that is, intraspecific competition must be greater than interspecific competition for both species. After a further bit of algebraic manipulation we see that a necessary (but not sufficient) condition for their coexistence is

$$1 - \alpha_{12}\alpha_{21} > 0. \tag{11}$$

Equations 10a and 10b are the formal conditions for competitive coexistence, whereas equation 11 indicates the minimal condition on the competition coefficients themselves. Note that with the original mechanistic derivation of these

equations (equations 2–6), which was based on two consumer (predator) populations consuming a single resource, the mechanistic meaning of α_{ij} is

$$\alpha_{ij} = \frac{a_j}{a_i},$$

whence equation 11 gives us

$$1 - \frac{a_j}{a_i} \frac{a_i}{a_j} = 0,$$

meaning that the inequality in equation 11 cannot be satisfied (under the assumption of a single resource). Thus it is not possible for two species to coexist on a single resource, something that Lotka noted in 1926, and this is effectively a proof of the competitive exclusion principle.

The competitive exclusion principle has been an important point of departure for a great deal of ecological theory, especially in community ecology. To this day, in highly elaborated forms, the principle still forms the basis of much thinking about how ecological communities are structured. However, the specific form of the classic Lotka–Volterra model (equations 6a and 6b) is too simple to reflect the complexities of the natural world. New models greatly complicate the entire field, and new ways of conceptualizing the competitive process have even challenged the underlying idea that no two species can occupy the same niche and coexist indefinitely. Some of these newer approaches are introduced in a later section of this chapter, but a comprehensive review is beyond the scope of this book.

The Competitive Production Principle: Applications of Competition Theory to Agriculture

Intercropping or, more generally, multiple cropping (Vandermeer 1989), is frequently observed in traditional forms of agriculture, especially in the tropics. The basic idea is that more than one species of crop is grown in the same field, such that the two species are competing with one another, a clear case of interspecific competition.

Exactly what are the ecological benefits thought to accrue from the practice of multiple cropping? The hypothesized ecological benefits have been divided into two categories (Vandermeer 1989), reduced competition (or the competitive production principle) and facilitation. In the case of the competitive production principle, which we discuss here, it is thought that two different species occupying the same space will use all the available resources more efficiently than a single species occupying that same space, thus corresponding to the coexistence criterion of the classic Lotka–Volterra competition equations, as described above.

The classical criterion for deciding whether an intercrop is better than its associated monocultures is the land equivalent ratio (LER) (which is equal to the relative yield total, RYT). If we wish to produce maize and beans in a field, would it be better to produce them together as an intercrop or to divide

the field into two parts and produce maize in one part and beans in the other? Presuming that the two crops compete with one another, we expect that the relative yield of each crop will be less than 1.0. That is, if we define the relative yield as P_i/M_i , where P_i is yield of crop i in the intercrop (P = polyculture) and M_i is its yield in monoculture, in both cases the yield expressed per unit area (kilograms per hectare, biomass per hectare, or some other relevant measure). Then the relative yield total is simply the sum of the two relative yields,

$$\text{RYT} = \text{LER} = P_1/M_1 + P_2/M_2.$$

The meaning of this measure is clear upon reflection. If the polyculture produces some specific amount in a fixed area (say an area of 1 hectare), how much area would be required to produce that same amount in two separate monocultures?

Thus $\text{LER} > 1.0$ is usually taken as the criterion as to whether the intercrop will perform better than the separate monocultures. It is certainly not the only possible criterion (Vandermeer 1989), and sometimes it can be misleading, but it is normally taken as the first step in analyzing an intercrop. If LER is less than 1.0, the two monocultures would be better, but if LER is greater than 1.0, yield would be better in polyculture.

Because $\text{LER} = 1.0$ is the critical value deciding whether an intercrop will be better than separate monocultures, we can write

$$1.0 = P_1/M_1 + P_2/M_2,$$

which can be rearranged as

$$P_1 = M_1 - [M_1/M_2]P_2. \quad (12)$$

If we plot equation 12 on a graph of P_1 versus P_2 , we see that it is a line connecting the two monocultures (figure 8.7). Furthermore, we can easily see that $\text{LER} > 1.0$ represents the area above the line and $\text{LER} < 1.0$ the area below the line. Depending on the strength of interspecific relative to intraspecific competition (see the explanation in figure 8.6), the system will be found either above or below the line connecting the two monocultures. The parallel between figures 8.6 and 8.7 is not accidental. The basic ecological process of competition is in operation in either case, and the question whether the intercrop will yield better than the two separate monocultures (see figure 8.7) is parallel to the question whether the two species will coexist (see figure 8.6). This parallelism between the classical ecological principle of competitive coexistence and polyculture advantage suggests that this form of polyculture advantage should be termed competitive production and that the parallel principle should be called the competitive production principle (Vandermeer 1989).

Resource Competition

The classic competition equations of Lotka and Volterra (equation set 6) may be classified as phenomenological in the sense that the process of competition is the phenomenon of concern, and it is modeled with a parameter that

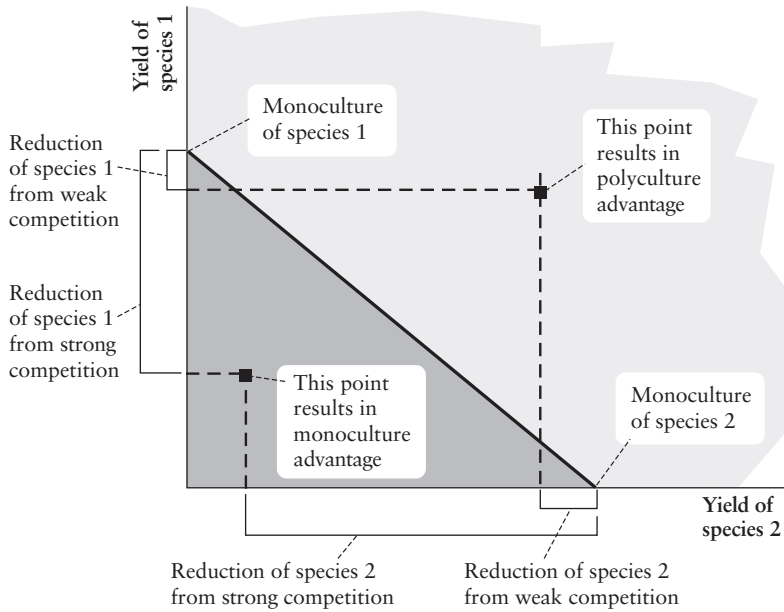


FIGURE 8.7. The graphic interpretation of the land equivalent ratio criterion (compare with figure 8.6). The dark shading indicates a monoculture advantage (below the critical line), and the light shading (above the critical line) indicates an intercrop advantage.

represents that phenomenon directly (the competition coefficients in equation set 6). Although that is a useful first step in approaching the topic, the alternative mechanistic approach is the basis for most of the more advanced treatments of the subject. With a mechanistic approach, ecologists focus on how the consumption behavior of the competing organisms translates into the observed phenomenon of competition. Thus, rather than beginning with competition coefficients, we begin with the competing populations consuming resources (or otherwise modifying the environments of one another) and ask how the process of consumption translates into patterns of competition.

EXERCISES

- 8.13** In a laboratory situation, two species of flour beetle have been cultured separately. Both populations approach an apparent carrying capacity asymptotically: that of the first is 200, that of the second 400. Additionally, both have been cultured together. The population densities they reach in mixed culture are 136 for the first and 120 for the second. Assuming that this competitive process is reasonably modeled by the classic Lotka–Volterra competition coefficients, what are the values of α_{12} and α_{21} ?
- 8.14** Flour beetles eat flour. Thus in exercise 8.13 we could theoretically estimate the rate of consumption of each of the two species by comparing their feeding rates. Recall the definition of the competition coefficient from the mechanistic derivation that started this chapter. If species 1 consumes at a rate of 0.25 grams per minute and species 2 at a

rate of 0.47 grams per minute, what might you estimate are the values of α_{12} and α_{21} ? Might this be a sensible way to measure relative competition for a single resource? How would the rates have to be arranged so as to ensure competitive coexistence?

The Lotka–Volterra competition equations are a phenomenological approach to questions about the population dynamics consequences of competition; the process of competition is described by the rather abstract parameters of competition coefficients and carrying capacity. As such, these equations have limited ability to explore questions about how the outcomes of competition might change in environments that differ in, say, availability of different resources or among species that differ in particular traits related to resource use. To address these kinds of questions, we now shift to a more mechanistic approach, that is, we explicitly model the dynamics of the resources for which competition occurs and allow species to interact solely through their consumption of shared resources. Mechanistic approaches to competition are most frequently based on the Lotka–Volterra predator–prey equations in one form or another, an idea we introduced in equations 2–4.

The first significant mechanistic approach to competition was that of Lotka (1926), whose original derivation of the competition equations (equations 2–6) was actually based on the mechanistic ideas of two populations consuming resources. However, perhaps the most historically significant mechanistic approach was that of MacArthur (1972), who had the insight to couple two predator–prey equations together to represent competition. Following the same development that led to equation set 6, we begin with the resource equation,

$$\frac{dR}{dt} = rR \left[\left(\frac{k - R}{k} \right) - a_1 N_1 - a_2 N_2 \right],$$

which has the equilibrium value

$$R^* = k[1 - a_1 N_1 - a_2 N_2],$$

where a_1 and a_2 are the consumption rates of the two consumer species (N_1 and N_2). Now suppose that there are two distinct resources, which leads to

$$R_1^* = k_1[1 - a_{11}N_1 - a_{12}N_2] \quad (13a)$$

and

$$R_2^* = k_2[1 - a_{22}N_2 - a_{21}N_1], \quad (13b)$$

where a_{ij} is the consumption of resource i by consumer j . We can write the consumer equations as

$$\frac{dN_1}{dt} = N_1(b_{11}R_1 + b_{12}R_2 - m_1)$$

and

$$\frac{dN_2}{dt} = N_2(b_{21}R_1 + b_{22}R_2 - m_2),$$

which are the same as equations 2a and 2b except that there are two resources rather than one and the relative conversion of a unit of resource i by consumer j is b_{ji} . They are also equivalent to the basic predator equations (see chapter 6), where R represents the prey (here we call it the resource). As earlier in this chapter, we presume that the resource dynamics are very fast relative to the consumer dynamics, such that we can use the equilibrium values of the resources in the consumer equations.

We then substitute $R_1^* + R_2^*$ for R , which gives us

$$\frac{dN_1}{dt} = [b_{11}k_1(1 - a_{11}N_1 - a_{12}N_2) + b_{12}k_2(1 - a_{21}N_1 - a_{22}N_2) - m_1]N_1$$

and

$$\frac{dN_2}{dt} = [b_{21}k_2(1 - a_{22}N_2 - a_{21}N_1) + b_{22}k_1(1 - a_{12}N_2 - a_{11}N_1) - m_2]N_2,$$

which can be expanded to

$$\frac{dN_1}{dt} = N_1[b_{11}k_1 + b_{12}k_2 - m_1 - (b_{11}k_1a_{11} + b_{12}k_2a_{21})N_1 - (b_{11}k_1a_{12} + b_{12}k_2a_{22})N_2]$$

and a similar equation for N_2 . Multiplying and dividing by $b_{11}k_1 + b_{12}k_2 - m_1$ gives us

$$\frac{dN_1}{dt} = (b_{11}k_1 + b_{12}k_2 - m_1)N_1 \left[\frac{b_{11}k_1 + b_{12}k_2 - m_1 - (b_{11}k_1a_{11} + b_{12}k_2a_{21})N_1 - (b_{11}k_1a_{12} + b_{12}k_2a_{22})N_2}{b_{11}k_1 + b_{12}k_2 - m_1} \right].$$

Finally, multiplying and dividing by $b_{11}k_1a_{11} + b_{12}k_2a_{21}$, we obtain

$$\frac{dN_1}{dt} = (b_{11}k_1 + b_{12}k_2 - m_1)N_1 \left[\frac{\frac{(b_{11}k_1 + b_{12}k_2 - m_1)}{(b_{11}k_1a_{11} + b_{12}k_2a_{21})} - N_1 - \frac{(b_{11}k_1a_{12} + b_{12}k_2a_{22})}{(b_{11}k_1a_{11} + b_{12}k_2a_{21})}N_2}{\frac{(b_{11}k_1 + b_{12}k_2 - m_1)}{(b_{11}k_1a_{11} + b_{12}k_2a_{21})}} \right].$$

Although this equation looks intimidating, its overall structure should be familiar (stare at it for a moment). Just make the following simple substitutions:

$$r_1 = b_{11}k_1 + b_{12}k_2 - m_1, \quad (14a)$$

$$K_1 = \frac{b_{11}k_1 + b_{12}k_2 - m_1}{b_{11}k_1a_{11} + b_{12}k_2a_{21}}, \quad (14b)$$

and

$$\alpha_{12} = \frac{b_{11}k_1a_{12} + b_{12}k_2a_{22}}{b_{11}k_1a_{11} + b_{12}k_2a_{21}}, \quad (14c)$$

and the equation becomes

$$\frac{dN_1}{dt} = r_1 N_1 \left[\frac{K_1 - N_1 - \alpha_{12} N_2}{K_1} \right].$$

Repeating the above derivation for N_2 , we get

$$\frac{dN_2}{dt} = r_2 N_2 \left[\frac{K_2 - N_2 - \alpha_{21} N_1}{K_2} \right].$$

The above two equations are identical to the Lotka–Volterra competition equations we derived above through a different logic. The beauty of this form, which we owe to Robert MacArthur (1972), is that we can easily see the meanings of the various parameters in a mechanistic sense, which is to say, in terms of the parameters of the underlying resource dynamics. Thus each of the classic parameters, carrying capacity (K), intrinsic rate of natural increase (r), and competition coefficient (α), can be seen as a combination of the parameters of the full dynamic resource model.

We can also use a graphical approach to derive the conditions for the coexistence of competitors using the same resources, as well as the traits that confer superior competitive ability when competitive exclusion occurs. The graphic approach was originally developed for a single resource and extended by Tilman (1982) to a second resource. This graphical approach also allows us to add a few more pieces of reality to the simple equations we have already developed. The key point is to recall that a competitive equilibrium in a mechanistic, or resource-based, model must have both the resources and the consumers at equilibrium. As with the development of the equations in the previous section, we will start with how a single species uses a single resource, add a second species competing for the same resource, and then add a second resource.

EXERCISES

- 8.15** Recall from chapter 6 that the basic exponential equation can be written with the birth and death process explicitly stated, as follows:

$$\frac{dN}{dt} = bN - mN.$$

When modeling a dynamic population whose birth rate is a diminishing-returns increase as a function of a resource, R , what will this equation look like, and what will a graph of $b(R)N$ and mN separately, as a function of R , look like? (Let the birth constant be 1, the functional response term [i.e., the term in the divisor that multiplies the resource] 5, and the death rate 0.1.)

- 8.16** Suppose that the population growth rate of a consumer is

$$\frac{dC}{dt} = \frac{aRC}{1 + bR} - mC,$$

where $aRC/(1 + bR)$ is the birth rate and mC is the death rate. For a particular value of C (say, 10), plot the birth rate and the death rate as a function of R for $a = 1$, $b = 5$,

and $m = 0.1$. Will C be increasing or decreasing when $R = 0.7$? Will C be increasing or decreasing when $R = 0.1$? What is the value of R at which C will be neither increasing nor decreasing?

- 8.17** Repeat exercise 8.16 but for $C = 15$. Shade the area where C will be increasing using a different color than that you use for the area where it will be decreasing. What is the equilibrium value for R ? Now change C to 20, and plot the results on the same graph. What is the equilibrium value for R now?

For mathematical simplicity, the previous derivation assumed a linear relationship between population growth rate and resources. However, in reality, as a resource becomes more available, it is increasingly likely that a different resource will become limiting. Therefore, growth as a function of a single resource will eventually asymptote (figure 8.8A). The graphical model incorporates this shape, although this doesn't change any fundamental results. Note that the asymptote in figure 8.8A is for maximum per capita birth rate and not for population size (as in the typical graph of logistic growth, such as that in figure 1.7 or in figure 8.8B). Population size is at equilibrium when the birth rate is equal to the mortality rate. In figure 8.8A, the mortality rate is shown as independent of resource availability, although this could be easily modified. The critical message of figure 8.8 is that when the consumer population is at equilibrium, the resource will also be at equilibrium; this amount is called R^* ("R star"). To see this, note on the graph that if R exceeds R^* , the birth rate (the solid line) exceeds the death rate, so the population size of the consumer increases. A larger population will take up more resources, so fewer resources will be available, that is, R will decline back toward R^* . In contrast, if R is less than R^* , the death rate is greater than the birth rate. Therefore, the population will decline in size, so resource availability will increase back toward R^* .

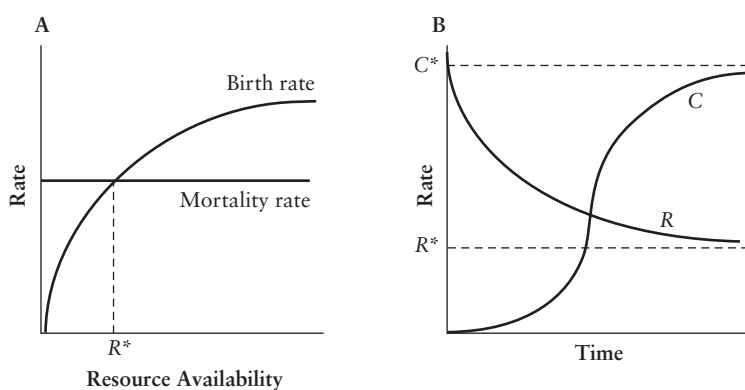


FIGURE 8.8. (A) Relationship between birth and mortality rates and resource availability, illustrating the meaning of R^* . (B) Growth patterns of the consumer population (C) and the resource population (R) over time.

Figure 8.8B shows the dynamics of the consumer–resource interaction. As C increases over time to an equilibrium (the asymptote in figure 8.8B), R decreases from some initially high value (where no individuals are taking it up) toward its equilibrium value at R^* . Where $C = C^*$, R is also constant at R^* . Note that R^* is a dynamic equilibrium just like C^* . At C^* , individuals are still being born and dying, although the population size stays constant. Similarly, at R^* , R is still being supplied at some rate (called the supply rate), but that supply rate is equal to the rate of consumption by the consumers.

Now add a second species competing for the same resource, for example, species C_B , with a lower R_B^* and a lower maximum potential birth rate than species C_A (figure 8.9A). Is coexistence possible? Or which species will be competitively superior, and which will be excluded? Here we present the graphic solution to these questions, but it is important to note that it has also been solved analytically by formulating equations for the birth and death functions in figure 8.9. Suppose that initially R is quite high, above the R^* of both species C_A and C_B . In this case, both populations are growing (with birth rate greater than death rate), and resource availability, R , will decline until it is below the R^* of species A. At this point, the death rate exceeds the birth rate for species C_A , and its population will decline. However, because species C_B is still growing, it will continue to deplete the resource even further, until R declines to the R^* of species C_B . This leads to the important and very general conclusion that under competition for a single limiting resource, the species with the lower R^* will always competitively exclude the species with the higher R^* (figure 8.9A). In other words, a species with a lower R^* can drive the resource to levels below that needed for inferior competitors to survive. As in the phenomenological Lotka–Volterra model or the simple resource competition model discussed above, two species cannot coexist on a single resource.

The dynamics of competition of two species for a single resource are shown in figure 8.9B. Note that the poorer competitor at equilibrium, spe-

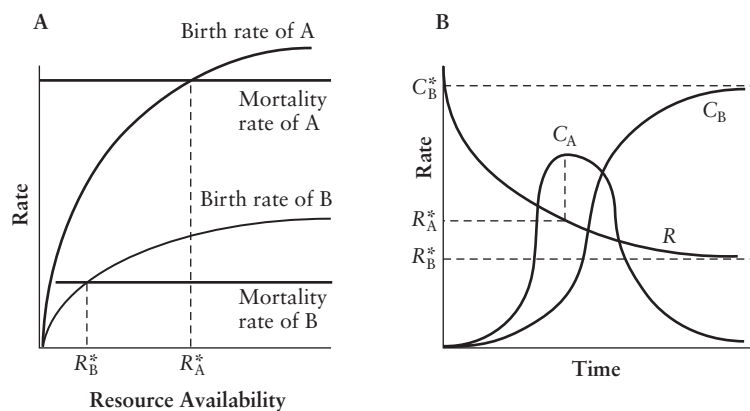


FIGURE 8.9. (A) Resource-dependent growth and mortality rates for two species using the same resource. (B) Time course expected for both species and the resource.

cies A, with the higher R^* , initially increases faster because it has a higher maximum birth rate when resources are high (figure 8.9A), even though it is excluded at equilibrium when resources have become scarce. This result has important implications for successional trends, as well as suggesting caution in the extrapolation of results from short-term competition experiments to the long-term outcome of competition.

What happens when a second resource is also potentially limiting? Is coexistence possible? Under what conditions? Tilman (1982) developed a graphical extension of the two species–one resource model to answer these questions, focusing on the equilibrial outcome. In figure 8.10 we plot the availability of resource 1 versus resource 2 and depict the combinations of the two resources that result in the consumer population’s being at equilibrium, or the zero net growth isocline (ZNGI) for the consumer. If the resources are substitutable, as they are for most animals, greater availability of one resource lowers the requirement for another resource so that the ZNGI is a straight line with a negative slope (figure 8.10A). However, if resources are essential, increases in only one resource do not change the equilibrium requirement for a second resource. This requirement is the R^* for each resource, so the ZNGI for essential resources is right angled, with the elbow at the R^* for both resources (figure 8.10B). Tilman analyzed a number of different kinds of resources in his 1982 monograph, but the model has been best developed for essential resources, and this is what is presented below. Essential resources are probably closest to reality for plants, which all require the same few resources: mineral nutrients such as nitrogen, phosphorus, potassium, calcium, and so on, as well as water and light (carbon dioxide and oxygen are also essential resources but are rarely limiting in terrestrial systems). Completely essential resources would mean that additional light, for example, could not compensate for less nitrogen. This is probably not strictly the case for plants; for example, more light could mean that more carbon is assimilated, and therefore more roots could be constructed to forage for nitrogen. This limited substitutability would result in a smoothing

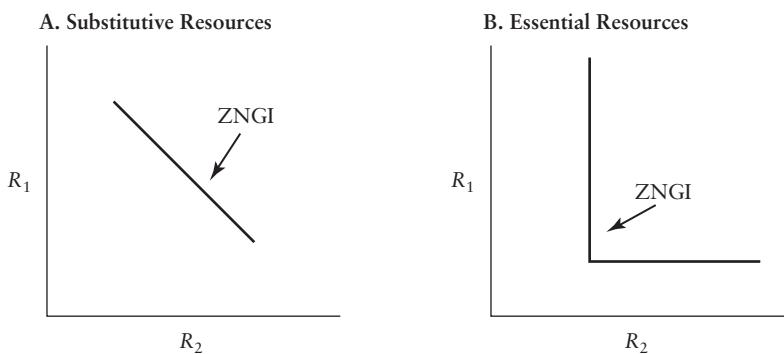


FIGURE 8.10. Zero net growth isoclines (ZNGI) for substitutive and essential resources.

of the right angle of the ZNGI but would not change any of the conclusions developed below.

A given point on a ZNGI (all points at C^*) is also at an R^* if the consumption rate is equal to the supply rate (u). Although the real consumption rate of any plant is a complex phenomenon, it can be greatly simplified if we assume that the consumer takes up the two resources in exactly the ratio in which they are needed. This would be, in a sense, an optimal strategy for a plant, because any additional amount of R_1 , for example, could not be used to increase the plant's growth rate because insufficient R_2 would be taken up. Therefore, the costs incurred in taking up the excess or "luxury" amount of R_1 would be wasted. Because the ratio of the two resources required is the ratio of their R^* s, we can simply draw a vector from the elbow of the ZNGI (where both $R_1 = R_1^*$ and $R_2 = R_2^*$) toward the origin to indicate consumption. The slope of this vector indicates the ratio of the two R^* s consumed, and its length indicates the amount of each resource that is consumed. At the elbow, the ratio required is also the ratio available. However, if plants are truly optimal foragers, they should consume the resources in the ratio required regardless of availability. Therefore, the same consumption vector should apply anywhere on the ZNGI (or indeed on the entire phase plane in figure 8.11A).

The supply vector can also be constructed graphically given a simple assumption. The supply vector is the rate at which the resource is transferred from unavailable to available forms. For example, the nitrogen found in an ecosystem is unavailable when tied up in organic matter, whether in living plants and animals or in litter and soil. Forms available to plants are usually mineral nitrogen, either nitrate or ammonium. The final step to available forms is called mineralization and is facilitated by microorganisms. If we make the simple assumption that this supply rate is proportional to the proportion of the resource that is unavailable, we have the simple function $u_j = a_j (S_j - R_j)$, where S_j is the total resource in the habitat (called the supply point), R_j is the available amount of resource j , a_j is a rate constant, and u_j is the supply rate. This assumption allows a simple graphical trick: starting from any combination of R_1 and R_2 on the phase plane, the supply vector will always point to the supply point (figure 8.11B). If, for example, most of resource 1

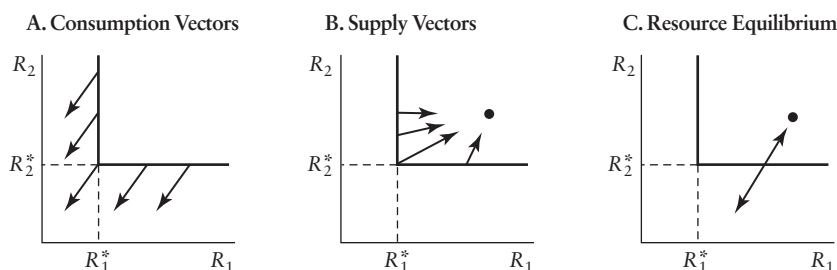


FIGURE 8.11. Consumption and supply vectors and resource equilibrium for the case of two essential resources.

is already in available form (high R_1 relative to S_1), little new resource can be supplied, and much more of resource 2 will be supplied than of resource 1. The length of the supply vector u is also determined by the supply point. When the total supply of a resource is much greater than the amount available, the supply rate will be higher (the vector will be longer) than when most of the resource is already in available form.

Combining the consumption and supply vectors, you can see that whether a given point on a ZNGI is both a consumer and a resource equilibrium point depends on the supply point (Figure 8.11C). For the resource to be at equilibrium, the slope and length of the supply and consumption vectors need to match exactly. Thus for any given supply point, assuming a constant and optimal consumption vector, there is a single combination of resource availabilities on the ZNGI that is an equilibrium point for both consumers and resources.

The final step in addressing questions about the outcome of competition of two species competing for two resources is to add a second species by adding a second ZNGI (figure 8.12). There are four possible ways of combining two ZNGIs (see figure 8.12). In cases 1 and 2, the two ZNGIs do not cross, and therefore there is no two-species equilibrium point and no coexistence. In case

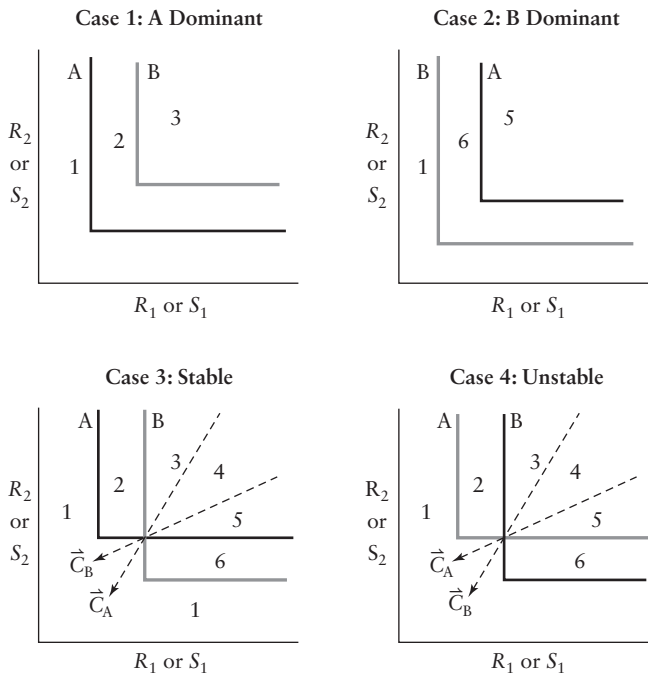


FIGURE 8.12. The four distinct cases of resource competition. Note that the slopes of the consumption vectors in case 3 indicate that both species are optimal foragers, whereas both species in case 4 are not. The numbers 1–6 designate habitats with different resource supply points (see text).

1, the ZNGI for species B lies inside that of species A, that is, species A has a lower R^* for both resource 1 and resource 2. If the supply point is below both ZNGIs, neither species can survive, regardless of competition. If the supply point is between the ZNGIs, the resources are too scarce for species B to survive, regardless of competition, but sufficient for species A. If the supply point is above both ZNGIs, species A will eventually deplete both resources to levels lower than the minimum required for species B to balance its mortality rate with its birth rate. Thus, regardless of supply point, species A will always competitively exclude species B. Case 2 is identical except that the two species are reversed.

In contrast, cases 3 and 4 both have the ZNGIs crossing and therefore a combination of availability of resources 1 and 2 for which both consumers have a stable population size. For the ZNGIs to cross, it is apparent that a trade-off in ability to compete for the two resources is necessary: the better competitor for resource 1 (lower R_1^*) must be an inferior competitor for resource 2 (higher R_2^*) and vice versa. This is the first requirement for coexistence, but remember that stability in a resource competition model also requires that the resources be at equilibrium.

To determine whether both resources are at equilibrium at the intersection of the two ZNGIs, we need to construct the joint consumption vector and then find the supply vector that exactly balances this consumption. The exact consumption vector depends on the relative abundance of the two species, but we know that it must lie between the consumption vector of a monoculture of species A and that of a monoculture of species B. In case 3, if we extend these backward to find supply vectors that balance these consumption vectors, we enclose a triangular area of possible supply points for which coexistence is possible (habitat 4 in case 3 of figure 8.12). What if the supply point is not in this “coexistence region”? In habitat 1, the resources are insufficient for both species, regardless of competition, while resource 1 is insufficient for species B in habitat 2, and resource 2 is insufficient for species A in habitat 6. Habitats 3 and 5 are the most interesting because, in the absence of competition, both species could persist if there were supply points in these regions. By examining what happens in these regions, we also gain some clues about exactly why coexistence is possible in habitat 4. Suppose that we have a supply point at X in habitat 3 (figure 8.13). If only species A or B were present, their equilibrium would be at point A_x or B_x , respectively. Therefore, both species are limited by resource 1 (on the vertical part of their ZNGIs), and because species A is a better competitor for resource 1, it will drive resource 1 to levels below the minimum required for species B. The opposite happens if the supply point is in habitat 5, that is, at point Y. In this case, both species are limited by resource 2 and the better competitor for resource 2, species B, competitively excludes species A. Now try the same exercise with the supply point Z in habitat 4, where the two species do coexist. In this case, species B is limited by resource 1, but species A is limited by resource 2.

This derivation gives us two important criteria for stable coexistence. First, the two species must be limited by different resources. Second, each species

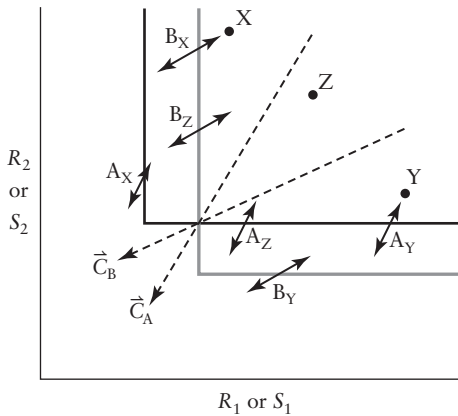


FIGURE 8.13. Equilibrial consumption and supply vectors in monocultures for two optimally foraging competitors with different supply points.

must be limited by the resource for which it is a poorer competitor. Species A is a better competitor for resource 1 but is limited by resource 2 in the habitats in which coexistence is possible. Species B is a better competitor for resource 2 but is limited by resource 1 in the habitats of coexistence. An intuitive explanation of this second criterion is that if a species were limited by the resource for which it is a better competitor, it would reduce the resource to its R^* , which, by definition, would be below that of the poorer competitors, and therefore no coexistence would be possible. Because optimal foragers always consume relatively more of the resource for which they are poorer competitors (i.e., for which they have a greater requirement), it turns out that as long as both species are optimal foragers, these two criteria will always hold. Therefore, as long as there is a trade-off in competitive ability among species for different resources, the two additional criteria for coexistence developed here should always hold for some possible supply points. Case 4 shows an example identical to that of case 3 except that neither species is an optimal forager and each species consumes relatively more of the resource for which it is a superior competitor. In this case, despite the existence of a two-species equilibrium point, the equilibrium is not stable.

To summarize, this graphical model of resource competition comes to essentially the same conclusions as the simpler phenomenological Lotka–Volterra competition models—that species must differ in resource use to coexist—but gives us a much more detailed understanding of the mechanisms of interaction and coexistence and how the interaction depends on environmental conditions. For example, one simple insight from this model provides a potent critique of a common interpretation of field data on distribution patterns. Suppose you saw that the abundance of a particular plant species increased with the availability of nitrogen in the soil. Your intuitive interpretation or hypothesis based on this correlation would be that nitrogen must be a limiting resource for that species and that the higher level of nitrogen is responsible for the greater plant abundance. However, the resource competition model suggests that exactly the opposite interpretation is correct! If a

resource is limiting (and a population is at equilibrium—an important caveat for this alternative interpretation), the population should drive the availability of that resource down to the same R^* regardless of the total supply of that resource in the environment (i.e., the supply point). Therefore, a positive correlation between population size and any resource should actually be regarded as evidence that that resource is not limiting because it is not being consumed down to its R^* and the positive correlation must be due to some other factor (e.g., the supply rates of different resources—a nonlimiting and a limiting one—may be correlated along a gradient because changes in moisture and temperature allow microbes to process both resources faster).

Facilitation and Mutualism

The situation in which one population has a beneficial effect on a second population may be referred to as facilitation, a phenomenon that is becoming more frequently recognized in population biology (Bruno et al. 2003). When two populations simultaneously have beneficial effects on one another, the phenomenon is referred to as mutualism. Mutualisms are arguably the most important interaction in nature if we consider all their varied forms. Consider the ubiquitous relationship between mycorrhizal fungi and vascular plants, the association of *Rhizobium* bacteria with leguminous plants, and the diverse community of bacteria in all mammalian guts, including ours. Or plants and their insect pollinators and avian dispersers. Or fish and their cleaning shrimp. Or ants and their homopteran partners. There are many.

Despite the ubiquitousness of mutualisms, no central and standard body of theory of their population dynamics has emerged in the way that it has for predator–prey systems or for competition, although the evolutionary ecology of mutualisms has been extremely well studied (Bronstein et al. 2006). Some have argued that this is because of the much greater diversity of mechanisms involved in mutualisms than in the consumer–resource interactions that lie at the heart of both predator–prey and competitive interactions. Nevertheless, at its most elementary level, some forms of mutualism might be understood through a simple extension of the Lotka–Volterra competition equations. This approach does not adequately address some of the more evident mutualisms described above, in which the interacting organisms have dynamics on very different time scales, such as that involving bacteria and human hosts. However, it is convenient to consider the sort of mutualism that exists between, for example, annual plants and their insect pollinators or other systems in which the time scales of population

change are comparable for the two populations. While this probably does not cover the majority of the mutualisms that actually exist in the world, it is useful as a point of departure for more complex approaches to the study of the dynamics of mutualistic interactions (e.g., Holland and DeAngelis 2010).

The equations that emerge are elementary, recalling the classic competition equations

$$\frac{dN_1}{dt} = r_1 N_1 \left[\frac{K_1 - N_1 + \alpha_{12} N_2}{K_1} \right] \quad (1a)$$

and

$$\frac{dN_2}{dt} = r_2 N_2 \left[\frac{K_2 - N_2 + \alpha_{21} N_1}{K_2} \right], \quad (1b)$$

with isoclines

$$N_1 = K_1 + \alpha_{12} N_2 \quad (2a)$$

and

$$N_1 = \frac{N_2}{\alpha_{21}} - \frac{K_2}{\alpha_{21}}. \quad (2b)$$

These are identical to the isoclines of the competition equations, with the exception that the constants, the competition coefficients, are now mutualism coefficients; that is, the isoclines have a positive slope (figure 9.1). Thus increasing the density of the mutualist of species i supports an increase in N_i above its carrying capacity (defined as equilibrium density when alone). Although densities of N_i itself above K decrease population growth, further increases of its mutualist increase population growth at a given density of conspecifics.

Interestingly, the case of mutualism, even though it is superficially similar to competition, generates some complications up front. Consider, first, the case of the size of the mutualism coefficients. In figure 9.1C and D we show the difference between large mutualism coefficients and small ones. Larger coefficients mean that the equilibrium density of species i increases more for a given increase in mutualist density. With small coefficients we see the generation of a stable node that is larger than either of the carrying capacities (figure 9.1C). But with large coefficients both populations increase indefinitely (figure 9.1D), much as does a single population under the assumption of exponential growth. Obviously populations cannot continue increasing forever, so some other controlling force will come into play eventually. But in terms of this simple approach to expanding the Lotka–Volterra equations to take into account mutualisms, the expectation is that with sufficiently large mutualism coefficients, the populations will escape any sort of control that might be implied by the two carrying capacities.

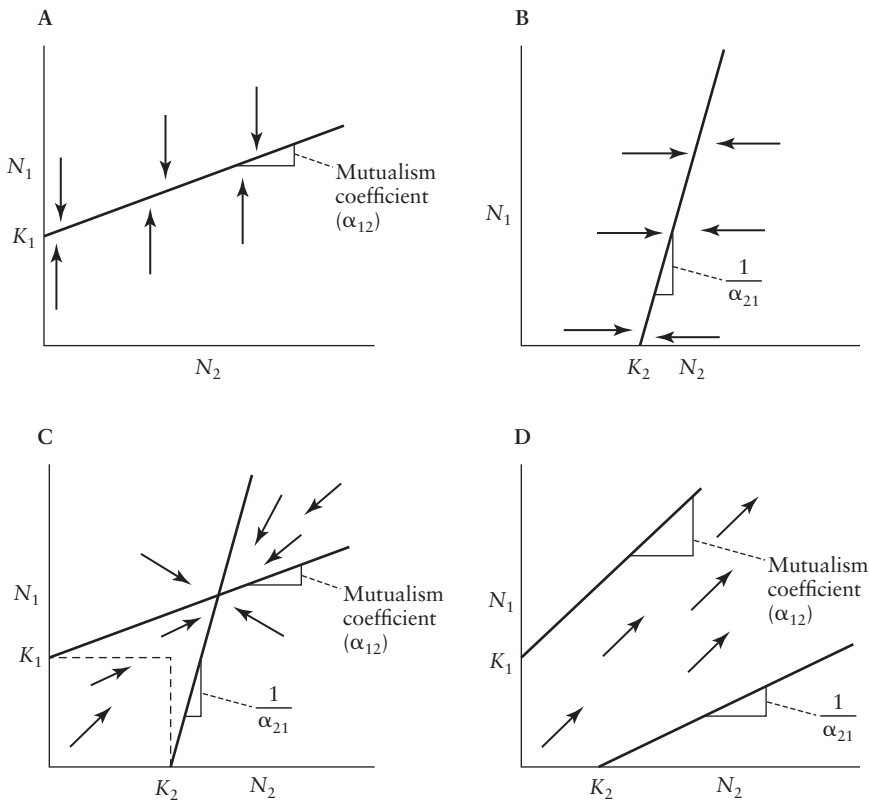


FIGURE 9.1. The basic dynamics of equations 1a and 1b, showing the two fundamental outcomes of facilitative mutualism. (A) Isocline for the first species (equation 2a) for weak mutualism. (B) Isocline for the second species (equation 2b) for weak mutualisms. (C) Isoclines for equations 2a and 2b plotted together with a summed vector field. (D) Isoclines for strong mutualism plotted together, illustrating the fundamental instability of strong mutualism.

EXERCISES

- 9.1** Derive the conditions for the α s that distinguish between a stable node and an unstable point at zero (i.e., the difference between figure 9.1C and D) for the basic mutualism equations.
- 9.2** Modify the basic mutualism equations to model the idea of an “obligate” mutualism (defined as a population that will tend toward zero if its mutualist partner does not exist).

Another aspect of the dynamics of mutualisms does not even arise in the analysis of competition and indeed would make no sense at all. Yet it is a key idea for many mutualisms, not just the sorts considered here. Many organisms are not just mutualists in that they benefit from the presence of another species; some actually require the presence of the other species to survive in

the first place. Thus they are in a sense obligated to be with the other species and are frequently referred to as obligate mutualists as opposed to their cousins that could exist without the other member of the mutualist pair but facultatively take advantage when their mutualist (or facilitator) is present, referred to as facultative mutualists. What we effectively model with equations 1a and 1b, assuming a real positive value of K , is facultative mutualism. Because in obligate mutualism, by definition, the populations go to zero in the absence of their mutualist partner, equations 1a and 1b can be approximated with the following:

$$\frac{dN_1}{dt} = r_1 N_1 [K_1 - N_1 + \alpha_{12} N_2] \quad (3a)$$

and

$$\frac{dN_2}{dt} = r_2 N_2 [K_2 - N_2 + \alpha_{21} N_1], \quad (3a)$$

and setting the carrying capacities equal to zero we have the isoclines

$$N_1 = \alpha_{12} N_2 \quad (4a)$$

and

$$N_1 = \frac{N_2}{\alpha_{21}}, \quad (4b)$$

which define the dynamics displayed in figure 9.2B. In figure 9.2A we illustrate more heuristically how the transformation from a facultative mutualist (N_1) to an obligate mutualist can be conceived as a gradual reduction of K_1 to zero, shown by the lightly dashed isoclines.

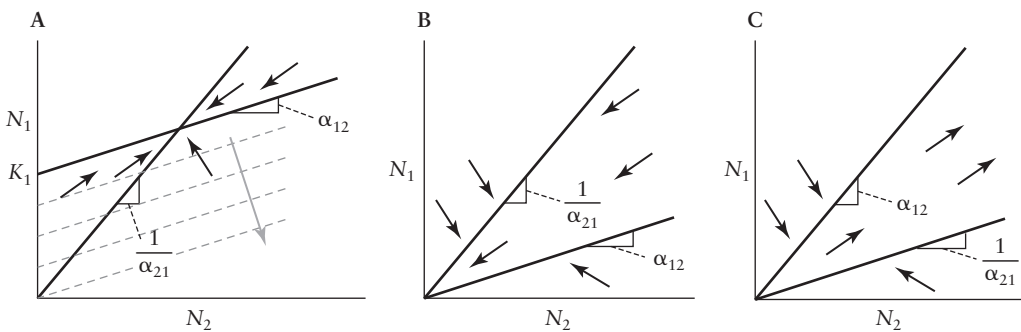


FIGURE 9.2. The various forms of isocline placement in the basic mutualism model. (A) Reducing the value of K_1 so as to reduce N_1 from a facultative mutualist to an obligate mutualist, as indicated by the lightly dashed isoclines and the lightly shaded arrow showing the sequence of changes in isoclines. (B) The isoclines of equations 4a and 4b with the determinant of the mutualist matrix greater than zero (i.e., showing weak mutualism). (C) The isoclines of equations 4a and 4b with the determinant of the mutualist matrix less than zero (i.e., showing strong mutualism).

EXERCISES

9.3 The detached coefficient matrix of equation set 3 at equilibrium is

$$\begin{vmatrix} 1 & \alpha_{12} \\ \alpha_{21} & 1 \end{vmatrix}.$$

Compute the determinant of this matrix (see Appendix A of chapter 2 if you need a reminder of matrix algebra), and relate it to the qualitative position of the isoclines in figure 9.2B (recall exercise 9.2).

9.4 What would reversing the sign of the determinant mean for the isocline graph?

Finally, in figure 9.2C we illustrate the alternative outcome of a basic obligate mutualism. So we see, in figure 9.2B and C, the two fundamental outcomes of strict obligate competition, a stable point at zero (i.e., both populations go extinct; figure 9.2B) and an unstable point at zero (i.e., both populations increase forever; figure 9.2C). The extinction outcome says, in biological terms, that the mutualism effect is not sufficiently strong to overcome the fact that the carrying capacity for the species is zero, whereas the ever-expanding outcome says the opposite. In a sense the relationship between the two mutualism coefficients changes the determinant of the mutualism matrix from a negative to a positive value (see exercises 9.3 and 9.4) and represents a kind of tipping point for a sudden change from the extinction of both species to the increase of both species without limit. Of course increasing without limit is ultimately unreasonable in the real world, and something else must happen eventually. However, from the point of view of understanding the mutualism per se, in strict obligate mutualism we have this result that either both populations go extinct or both increase forever.

One other feature of elementary mutualisms makes them uniquely different from competition and indeed shows how what we call strict obligate mutualism is actually a special case. It is probably rare in nature for an obligate mutualist species to be responsive to only a small number of its mutualist partners. So, for example, the population of annual plants that is dependent on a species of bee for pollination will probably die out if the bee population is too small to efficiently provide the pollination service. In the case of obligate mutualisms, the minimal population density of the mutualist can be incorporated into the basic equations (equations 3a and 3b, after setting $K = 0$) as follows:

$$\frac{dN_1}{dt} = r_1 N_1 [\alpha_{12}(N_2 - \phi_2) - N_1] \quad (5a)$$

and

$$\frac{dN_2}{dt} = r_2 N_2 [\alpha_{21}(N_1 - \phi_1) - N_2], \quad (5b)$$

where ϕ_i refers to the critical density of species i that will allow species j to increase. By observation (equations 5a and 5b) we see that if the population

density (N_i) is less than the critical population density (ϕ_i), the derivatives in equations 5a and 5b must always be less than zero (i.e., the population must decline). Thus the only way for either of the populations to increase is for N_i to be greater than ϕ_i .

It is possible to perform some seemingly arbitrary algebraic manipulations of equations 5a and 5b to make them look like equations 1a and 1b, but with an interesting twist. Let us define the variable $K'_j = \alpha_{ji}\phi_i$ and substitute into equations 5a and 5b, obtaining

$$\frac{dN_1}{dt} = r_1 N_1 \left[\alpha_{12} \left(N_2 - \frac{K'_1}{\alpha_{12}} \right) - N_1 \right]$$

and

$$\frac{dN_2}{dt} = r_2 N_2 \left[\alpha_{21} \left(N_1 - \frac{K'_2}{\alpha_{21}} \right) - N_2 \right].$$

With some more substitution and rearrangement, we find that

$$\frac{dN_1}{dt} = K'_1 r_1 N_1 \left[\frac{-K'_1 - N_1 + \alpha_{12} N_2}{K'_1} \right] \quad (6a)$$

and

$$\frac{dN_2}{dt} = K'_2 r_2 N_2 \left[\frac{-K'_2 - N_2 + \alpha_{21} N_1}{K'_2} \right], \quad (6b)$$

which has precisely the same form as the original mutualism equations, with the intrinsic rate of increase redefined and with the carrying capacity in the numerator negative. Obviously one cannot have a negative carrying capacity, but it is easy to see how the appearance of a negative carrying capacity in the equations is simply a consequence of adding the idea that there is a minimal number of individuals of the mutualist partner that can actually initiate the mutualism process and result in increasing population numbers.

Setting the derivatives in equation set 6 equal to zero and solving for N_1 , we have the isoclines

$$N_1 = -K'_1 + \alpha_{12} N_2 \quad (7a)$$

and

$$N_1 = \frac{K'_2}{\alpha_{21}} + \frac{N_2}{\alpha_{21}}, \quad (7b)$$

which are diagrammed, along with the vector field, in figure 9.3C and D. When mutualism is weak, the addition of the critical mutualist density (ϕ) makes no difference to the qualitative outcome, the extinction of both species. However, the addition of the critical mutualist density to the case of strong mutualism changes the system qualitatively from one in which both mutualists experience an ever-expanding population density to a situation of indeterminacy, with some initiation points resulting in extinction, others resulting in expansion.

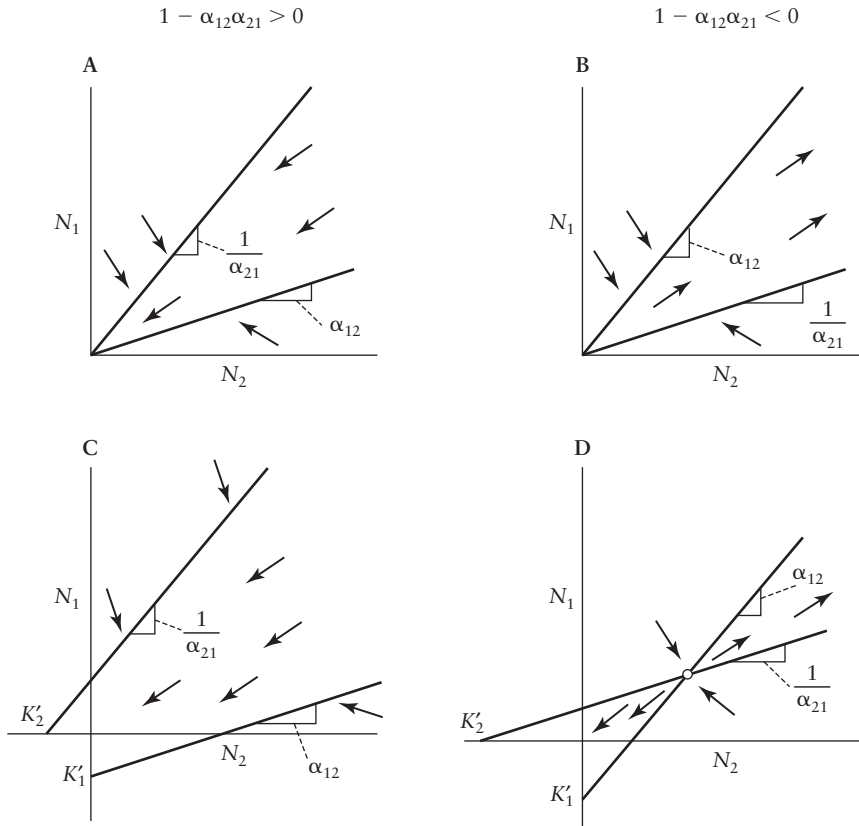


FIGURE 9.3. Complications with obligate mutualisms. (A) and (B). The same as figure 9.2B and C. (C) Adding the critical number of mutualists to the system of weak mutualism (note that the qualitative outcome is precisely the same as with a critical number zero (i.e., A and C have the same overall dynamic behavior)). (D) Adding the critical number of mutualists to the system of strong mutualism in which a qualitatively different result occurs (i.e., there is an unstable equilibrium point [a saddle point] separating two regimes, extinction of both species versus continual increase of both species).

All four possible cases of symmetrical mutualism are summarized in figure 9.4. Other cases in which the nature of the mutualism is not symmetrical (e.g., when one species is facultative and the other obligate) are easily derived from these basic four cases, as we discuss later (also see Vandermeer and Boucher 1978).

An evident problem with this simple development is that several basic arrangements (figure 9.4B and D) result in both populations' increasing in perpetuity. Much as in case of the original conundrum of bacteria growing exponentially forever, it is biologically unreasonable to have populations increasing forever. The most evident solution to this problem is to relax the assumption of linearity, which is most easily done using the framework we

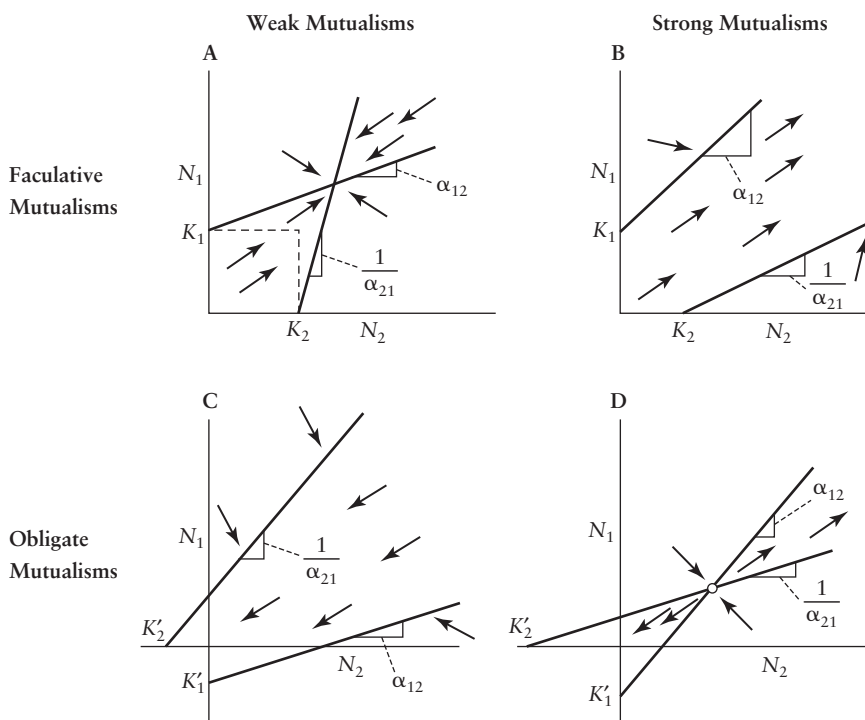


FIGURE 9.4. Summary of the qualitative results with symmetrical mutualisms (i.e., both obligate or facultative and both strong or weak).

used in chapter 6 with the graphic construction of isoclines (the framework of Noy-Meir 1975 and Rosenzweig and MacArthur 1963).

Recall the basic argument for a facultative mutualism. If species 1 exists alone, it will grow to its carrying capacity (figure 9.5A). Adding some small density of species 2 will permit species 1 to increase even if it is above its defined carrying capacity (figure 9.5B). The mechanisms that allow this are varied. For example, a self-pollinating plant could reproduce just fine without its mutualist pollinator, and its population would reach some carrying capacity. But if outbred seeds have greater fitness, when a number of pollinators arrive on the scene, the effective population size of that plant will increase due to the positive effect of the pollinator. Many other mechanisms are possible, and pollination is just a handy reference at this point. Some mechanism creates this facultative effect of species 2 on species 1, effectively creating a larger carrying capacity for species 1 than had existed before. And as the density of species 2 increases, the effective size at which species 1 can still increase its population will continue increasing (i.e., its effective carrying capacity will increase). However, the system must inevitably reach some point at which, whatever the detailed mechanism, further increases in species 2 will elicit less of the facilitative effect (figure 5D and E). Hundreds of pollinators are prob-

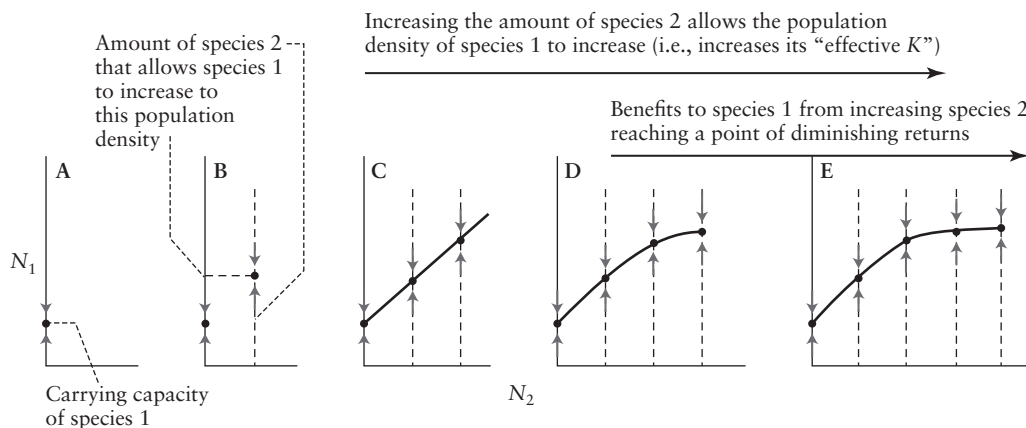


FIGURE 9.5. Transforming the isocline of equation 1a. (A) When the density of species 2 is zero, the population of species 1 will approach its carrying capacity. (B) Adding a small density to the population of species 2 (a facilitator of species 1) results in species 1 having positive population growth even when above its carrying capacity (effectively increasing the carrying capacity). (C) Adding yet more of species 2 allows species 1 to increase at yet larger densities. (D) and (E) Eventually a point must be reached at which additional density of the facilitator has minimal effect.

ably better than tens, but thousands probably add little to the net facilitative effect on the plant. The ultimate result is a simple nonlinear response, creating an isocline that exhibits diminishing returns.

In figure 9.6 we illustrate the result of adding such nonlinearities to the basic structure of the uncontrolled form in both facultative and obligate mutualisms.

EXERCISES

- 9.5 How would you modify equation set 1 to include the logic of figure 9.5 analytically? Hint: Recall the definition of a type II functional response from chapter 6.
- 9.6 Find the isoclines of the system derived in exercise 9.5, and plot them in the phase plane.
- 9.7 Repeat the exercise in figure 9.5, but give the N_2 isocline the saturating feature while N_1 remains linear.
- 9.8 Modify equation set 1 to include both a mutualistic coefficient and a competition coefficient, both of which are constant. What can you conclude about the overall behavior of the system?

Recalling the foundational equations of mutualism (equations 1a and 1b), the logic of adding the nonlinear component (see figure 9.5) can be analytically represented with a simple saturating function (Holland and DeAngelis 2010), as was done in the case of predator–prey interactions in chapter 6. Thus modifying equation set 1, we obtain

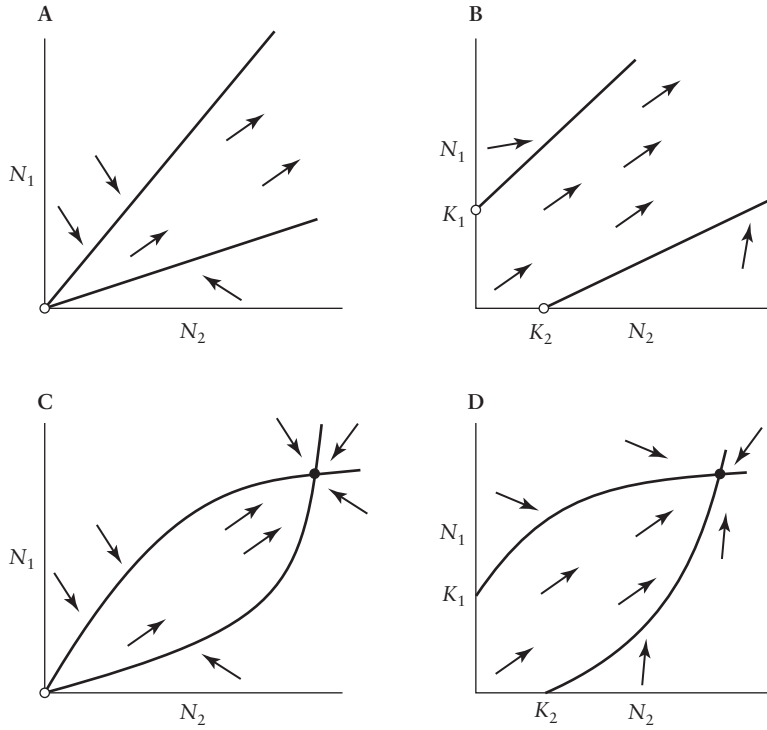


FIGURE 9.6. Dynamic consequences of adding nonlinearity in the facilitative effects of each of the mutualists for the case of high mutualism coefficients. (A) Obligate mutualism with constant mutualism coefficients (a repeat of figure 9.3B). (B) Facultative mutualism with constant mutualism coefficients (a repeat of figure 9.1B). (C) Modifying the isoclines of A with the diminishing-returns mechanism to produce nonlinear isoclines and a stable equilibrium. (D) Modifying the isoclines of B with the diminishing-returns mechanism to produce nonlinear isoclines and a stable equilibrium.

$$\frac{dN_1}{dt} = r_1 N_1 \left[\frac{K_1 - N_1 + \frac{\alpha_{12} N_2}{h_1 + N_2}}{K_1} \right] \quad (8a)$$

and

$$\frac{dN_2}{dt} = r_2 N_2 \left[\frac{K_2 - N_2 + \frac{\alpha_{21} N_1}{h_2 + N_1}}{K_2} \right], \quad (8b)$$

with isoclines

$$N_1 = K_1 + \frac{\alpha_{12} N_2}{h_1 + N_2} \quad (9a)$$

and

$$N_1 = \frac{h_2(K_2 - N_2)}{N_2 - K_2 - \alpha_{12}}, \quad (9b)$$

corresponding qualitatively to the situations in figure 9.6C and D.

Combining the insights of a nonlinear functional response for the mutualism coefficient along with the idea of a minimal critical size for mutualism (represented by a “negative” K'), we depict all possible qualitatively distinct forms of mutualism at this level of analysis in figure 9.7.

Further complications are involved when considering other aspects of mutualism. For example, Holland and DeAngelis (2010) take a consumer–resource approach, defining three distinct forms (bidirectional, unidirectional, and indirect), all of which are conveniently modeled by adding a negative consumption term to one or both of the basic equations (equation set 8). This complication generates yet further qualitatively distinct results, including three stable equilibria, a result that is certainly interesting, but beyond the scope of this book.

However, one particular form is of special interest in that it combines a couple of different basic population ecology processes already discussed in this text. In the particular case of mycorrhizae and plants, the action of the “mutualist” fungus ranges from parasitism to mutualism (Hoeksema et al. 2010). (A similar phenomenon occurs in many other mutualisms, such as those involving plants with specialist pollinators that also act as seed predators or herbivores [Holland et al. 2002].) Vannette and Hunter (2011) examined the mechanisms underlying the contingency of this interaction for a species of milkweed and one of its associated mycorrhizal fungi. The facilitative effect of the fungus is thought to aid in the protection of the plant; with the extra energy and nutrients supplied by the fungus, the plant is able to produce more of the chemical defenses needed to thwart herbivores. So the expectation is that with no mycorrhizal infection, the milkweed’s defenses against the herbivores will be low but will increase as infection increases, but only up to a point. As infection levels become very high, the cost of maintaining the fungus will become excessive, reducing the plant’s potential protection against herbivores (figure 9.8).

We can think of the mycorrhizal fungus as a “predator” stealing energy (carbon) from the milkweed, thus reducing its rate of biomass accumulation at the same time that it provides nutrients, thus increasing its potential rate of biomass accumulation when nutrients are limiting. The balance of these costs and benefits results in the nonmonotonic relationship shown in figure 9.8. The equations that we might write, then, using P as plant biomass and M as mycorrhizal biomass, are

$$\frac{dP}{dt} = r_P P \left(\frac{K_P - P}{K_P} \right) - \frac{aMP}{g + P}, \quad (10a)$$

and

$$\frac{dM}{dt} = \frac{aMP}{g + P} - mM. \quad (10b)$$

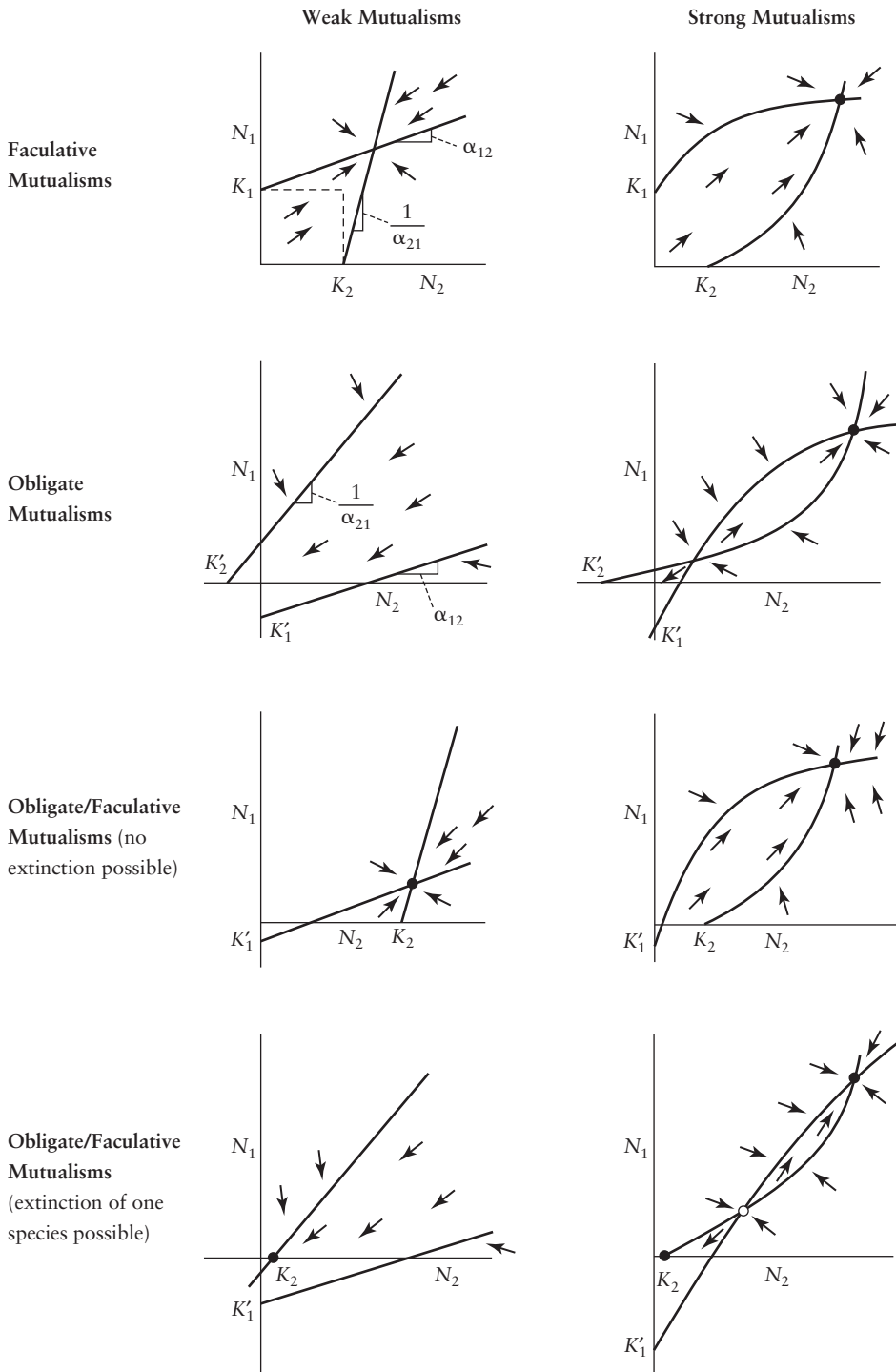


FIGURE 9.7. The eight qualitatively distinct cases of mutualistic interactions from the Lotka–Volterra extension, with nonlinear functional response added when necessary to prevent increasing densities of both species in perpetuity.

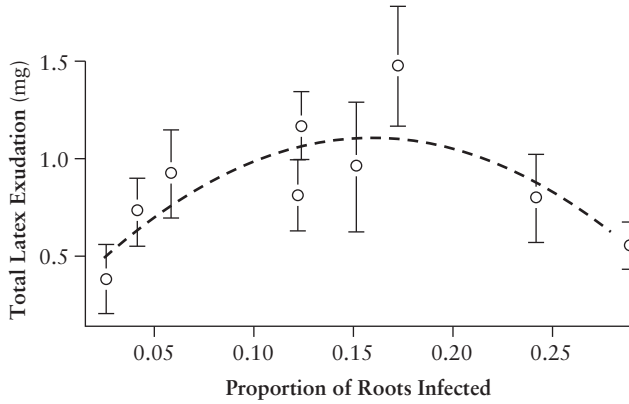


FIGURE 9.8. Production of latex (for defense against herbivores) as a function of the proportion of roots infected with mycorrhizal fungi. The dashed line is the best-fit quadratic function.

which are precisely the standard predator–prey equations (see chapter 6), with r_p the intrinsic rate of increase of plant biomass, K_p the carrying capacity of the plant, a the attack rate of the predator, g the handling time of the predator, and m the death rate of the predator (remember, in this form we are considering the mycorrhizal fungus a predator).

However, the mycorrhizal fungus is also a mutualist, which means that we should use the mutualism equations (e.g., equation set 1 as a first approximation), which means that we write

$$\frac{dP}{dt} = r_p P \left(\frac{K_p - P + \alpha_{PM} M}{K_p} \right) \quad (11a)$$

and

$$\frac{dM}{dt} = r_M M \left(\frac{K_M - M + \alpha_{MP} P}{K_M} \right). \quad (11b)$$

Now let's combine the two approaches. Let us presume that the mycorrhizal fungus is an obligate mutualist, that is, that it cannot live without consuming part of the milkweed's produce, which we can model by letting $r_M = 0$. Also, to make the analysis a bit simpler, let us suppose that the carrying capacity of the plant is 1.0 (i.e., all plant biomasses will be represented as some proportion of the maximum biomass) and that $r_p = 1$ (none of the dynamics change if r_p has different values). Putting equation sets 10 and 11 together, we have

$$\frac{dP}{dt} = P(1 - P + \alpha_{PM} M) - \frac{aMP}{g + P} \quad (12a)$$

and

$$\frac{dM}{dt} = \frac{aMP}{g + P} - mM \quad (12b)$$

as the final model. Setting derivatives equal to zero and rearranging, we compute the isoclines as

$$M = \frac{g + P(1 - g) - P^2}{1 - \alpha_{PM}(P + g)} \quad (13a)$$

and

$$P = \frac{gM}{a + m}. \quad (13b)$$

EXERCISES

- 9.9** Graph the isoclines for equation set 13 given the parameters $g = 0.25$, $\alpha_{PM} = 0.00$ (i.e., there is no mutualism, just predation), $a = 0.08$, and $m = 0.055$. What do you conclude about the dynamics of the system (i.e., what kind of equilibrium do you expect)?
- 9.10** Repeat exercise 9 but with $\alpha_{PM} = 0.05$. What do you conclude about the dynamics of the system, and how did adding the mutualism affect them?
-

Finally, it is worth noting that a large body of research has recently been generated associated with complex communities of mutualists (e.g., Bascompte et al. 2006). Again, this material is well beyond the scope of this book.

What This Book Was About

The subject matter presented in this text is logically organized, at least in the minds of its authors. In figure 10.1 we illustrate our vision of how the chapters are related to one another.

At the base is the subject matter of chapter 1, the central ideas of exponential growth and density dependence (one example of which is the logistic

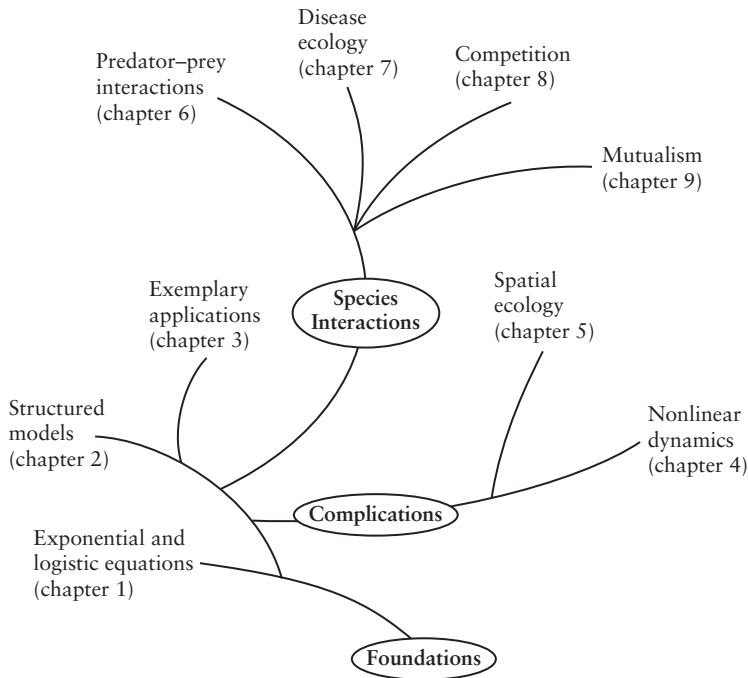


FIGURE 10.1. Diagrammatic representation of the subject matter of the text.

equation). From that base we branched out, extending that subject matter to the complications involved in classical structured models (chapter 2). Then we moved into some exemplary applications (chapter 3). The main branch also led to two generalized themes, complications and species interactions. We treated the complications as two distinct subjects. First, we gave an elementary summary of some of the exciting new areas that stem from the application of nonlinear dynamics to population ecology (chapter 4), and second, we delved into the complicated issue of spatial ecology (chapter 5). Finally, we presented four issues concerned with species interactions—predator–prey interactions (chapter 6), disease ecology (chapter 7), competition (chapter 8), and mutualism (chapter 9). Each of the tips of the branches on this tree (see figure 10.1) could be extended to more complicated analysis. This text is intended as a springboard to all of those more sophisticated analyses.

It is also useful to say something about what is not in this text. Part of the science of ecology stems directly from the material presented here. The basic question that drives much of ecological thought is that of species diversity—for example, why are there more than 200 species of trees in a hectare of tropical rain forest while there are only three in a hectare of boreal forest? This is, in its broadest outlines, the subject of community ecology. At least in principle, it should be possible to extrapolate from the two species in competition of chapter 8 to the multiple species in competition in a community of grass species in a savannah. Indeed, an analysis of the community matrix, introduced for a two-species system in chapter 8 but analyzed in the literature for a multiple-species system, forms the basis of much classical community analysis (e.g., Levins 1968; May 1973). An alternative but related approach begins with the predation equations of chapter 6 and constructs complicated food webs based on their structure (e.g., Baskerville et al. 2011; Hastings and Powell 1991; Pascual and Dunne 2006).

But community ecology is only the most obvious extension of the material in this book. Chapter 7 introduces the analysis of microparasites, which requires a distinct break with the population dynamics of exponential and logistic equations and uses the metapopulation (with individual animals or plants as the habitat patches and the disease organism as the occupant of those patches) as the underlying theoretical framework. This is the subject of epidemiology or, more generally, disease ecology, which has its own literature, to say nothing of its own departments in most major universities. And here the literature is truly staggering, with a history predating most of the rest of population ecology (indeed, epidemiological studies of schistosomiasis alone far outnumber and predate most other approaches in population ecology, with one reference [Warren and Newill 1967] listing more than 10,000 papers on the subject). Yet we think it not excessively hubristic to claim population ecology as the fundamental organizing principle of epidemiology. Indeed, one of the most highly respected treatises on human disease was written by two ecologists (Anderson and May 1991). Furthermore, a plausible argument can be made that an ecological approach to human disease is the wave of the future (or should be).

A third field that emerges from population ecology is the study of spatial patterns in biological populations and communities. This field, which has burgeoned after the popularization of GIS technology, is based on the elementary introductory material supplied in chapter 5. That chapter introduces not only the idea of clumping in space but also the idea of spatial scale, which has become exceedingly important (e.g., see Pascual and Levin 1999; Pascual et al. 2001). Moreover, we used the fact of spatial clumping to introduce the idea of metapopulations, the subject of an enormous literature in recent years (summarized in Hanski and Gaggiotti 2004).

Mastery of the material in this book will provide students with the background to delve into most of the literature on these more advanced topics with a sound analytical base. However, there is a perhaps more important use for the material presented in this text. Many of the world's most pressing problems will be solved only through the application of this material. In much the same way that physics and chemistry supplied the underlying scientific basis for engineering during the late phases of the Industrial Revolution, and similar to the way in which molecular biology is now emerging as the underlying scientific basis for the development of new medical technologies, population ecology is the scientific base from which solutions to many of the world's most difficult problems will be found. Although we do not mean to denigrate the immediate problems facing us in the year 2013, and we would sincerely love to see peace in the Middle East, economic development in Central America, and a resolution to the ongoing conflicts in the Congo, we insist that there are even deeper and more difficult problems associated, either directly or indirectly, with ecology. Indeed, there will be no final peace in the Middle East until we solve the problem of access to water, an issue of sustainable resource use. Central America will not see economic development until there is an end to the destruction of its natural environment, a problem of ecology as much as of politics. And resolution of the conflict in Congo will not completely solve the problem of epidemic diseases currently ravaging sub-Saharan Africa. These are all problems of ecology, and they will not be solved without attention to fundamental principles, the most elementary of which are the subject of this book.

Examples of practical applications have been peppered throughout the text, sometimes contrived, sometimes real. For the problems most obviously associated with the principles of population ecology we could force a classification based on what we see as emerging communities of scholars in contemporary society. First, for the effective use of natural resources, from water in the Middle East to land in Africa to petroleum in the world, the set of principles of resource consumption is an essential foundation. Chapters 1 and 2 specifically provide the foundations upon which most of the literature on natural resource harvesting is based (e.g., Getz and Haight 1989).

Second, the current crisis of the loss of biodiversity is ultimately related to the material presented in chapters 5, 6, and 8. Determining what allows large numbers of species to coexist (Hubbell, 2001; Palmer 1994) and thus what needs to be modified to maintain all of those species requires a knowledge of the material in those chapters. Although it is obvious that most of this prob-

lem will not be resolved without substantial political change, it is also the case that even if we develop a utopian politics, human intervention will still threaten the world's biodiversity without a better understanding of what that biodiversity means in the first place.

Third, and somewhat more cryptically, there is a growing realization that the world is an uncertain place. The recent revolution in nonlinear dynamics and complex systems has heightened our appreciation of the principle of "the inevitability of surprise" in any ecological system (Levins, personal communication). Indeed, much of the current appreciation of the uncertainty of complex dynamic systems was propelled by the work of an ecologist (May 1974). We presented the material in chapter 4 in an attempt to introduce the student to the principles underlying this new field. Applications here are hard to predict, but practitioners are not averse to claiming grand breakthroughs in understanding. In application to the natural world, this field promises to change at least some of the way we view environmental problems (e.g., Hastings et al. 1993; Schaffer 1985; Sheffer 2009).

Fourth, there is an emerging field called ecosystems management. Probably most of the world's rural residents would be surprised to hear that we academics have discovered what they and their ancestors have been doing for so many years, but so goes the academy. Thus we are now concerned with the consequences of human activities for "ecosystem services." Here, too, the principles introduced in this text are essential. For example, the management of agroecosystems has, for many years, required the development and use of techniques for pest control, many of which have relied on some sort of predator or parasite. The theory presented in chapter 6 is the underlying basis for the sometimes very sophisticated analytical approaches applied to this problem (Hawkins and Cornell 1999).

Fifth, who can ignore the fact that new diseases are emerging yearly and old diseases are emerging as new problems? The disaster of AIDS in Africa is fundamentally a problem of population ecology, as is the emergence of the Ebola virus. Currently there appears to be a spread of malaria into the East African highlands, likely due to environmental factors, and the recent resurgence of pertussis in the United States is completely enigmatic thus far. To deal with these problems, in addition to the obvious public policy necessities, we need to more fully understand the epidemiology of each of these diseases. For this, the material presented in chapter 7 is critical.

These five sets of practical problems are intimately related to one another, and it would be foolish to suggest that the ecological principles involved in any way trump the political and economic barriers to their solution. However, each of them, and all of them together, contain problems that will not be solved, no matter how well the politics and economics are dealt with, without some attention to the underlying ecological principles. Introducing those principles was our purpose in writing this book. Students who master this material will be poised to delve into more sophisticated theoretical and empirical ecology as well as to understand some of the underlying basics needed to solve practical problems of the environment. We hope our presentation makes that mastery agreeable.

GLOSSARY

- attractor** An attractive region of the space defined by one or more variables. All trajectories not contained in that region will eventually wind up in the region. An attractor may be a point or a cycle that is an equilibrium and generates transients that return to the equilibrium state after perturbation. It may also be an attractive region that has no individual equilibrium points or cycles (a chaotic or strange attractor).
- basin of attraction** The collection of points that converge on a particular attractor.
- basin boundary collision** A type of bifurcation in which the edge of a basin of attraction intersects with the edge of a strange attractor.
- bifurcation diagram** A graph of the attractors of a system as a function of some parameter (the “bifurcation” parameter).
- bifurcation point** A point of structural instability in which a single equilibrium condition is split into two.
- carrying capacity** The maximum attainable size of a population, usually symbolized as K .
- chaos** Behavior of a system that is inherently unpredictable in the sense that two points of initiation that are extremely close together will generate trajectories that deviate from one another dramatically.
- competitive exclusion principle** The idea, derived from the simple Lotka–Volterra equations, that if interspecific competition between two species is sufficiently large, both species cannot coexist in the same environment.
- competitive production principle** The competitive exclusion principle applied to an agricultural situation in which two crop species will overyield if the competition between them is sufficiently slight.

- density dependence** The condition in which the rate at which a population increases or decreases is a function of its density (in contrast with density independence); often used interchangeably with the term *intraspecific competition* although not formally synonymous.
- density independence** The condition in which a population increases or decreases without relation to the density of the population (in contrast with density dependence).
- elasticity** The degree to which the population growth rate changes as a function of changes in the elements of the projection matrix, expressed as a proportion.
- equilibrium point** The value of a variable that does not change under the rules of a dynamic model. An equilibrium point may be stable (in which case it is commonly referred to as an attractor) or unstable (in which case it is commonly referred to as a repeller).
- Euler's constant** Approximately 2.7183, the base of natural logarithms, normally symbolized by a lowercase e .
- exponential function** Euler's constant raised to some value x , frequently symbolized by $\exp(x)$.
- exponential growth** A pattern of increase (or decrease) in a population that follows the exponential equation, either the integrated form (equation 10 in chapter 1) or the differential equation form (equation 8 in chapter 1).
- facultative mutualism** Mutualism in which one species can survive without its mutualist but performs better with it.
- functional response** In consumer–resource (predator–prey) equations, the function that stipulates how the per capita consumption rate (or predation rate) changes with changes in resource density.
- Gini coefficient** A measure of the inequality of a distribution of factors such as size or biomass.
- intraspecific competition** The competitive interaction among individuals in the same population.
- intrinsic rate of natural increase** The growth of a population under the theoretical state of extremely low population density, usually symbolized as r .
- isocline** For a dynamic model, the set of all points for which one of the variables does not change (in the context of a differential equation, the set of points that corresponds to the derivative set equal to zero).
- K selection** Selection for increased carrying capacity, usually symbolized as K , that occurs when a population is usually at high density, near carrying capacity, in contrast to r selection, which occurs

at low population density, far below carrying capacity. Often assumed to result in low reproduction potential and high survivorship.

limit cycle An oscillatory system that can be either stable (an oscillatory attractor) or unstable (an oscillatory repeller).

logistic population growth Population growth that appears qualitatively exponential at low population density but approaches an asymptote as the population becomes larger; population growth that follows the logistic equation (equation 17 in chapter 1).

metapopulation A population distributed in patches in which each of the patches is incapable of indefinitely maintaining a viable population alone (i.e., extinction probability > 0) but the population is maintained over the whole collection of patches because migrants from occupied patches continually reoccupy patches in which subpopulations have gone extinct.

obligate mutualism Mutualism in which one species is unable to survive without its mutualist.

one-dimensional map A function that projects a single variable through discrete time, for example, $N_{t+1} = f(N_t)$.

paradox of enrichment The tendency of a predator–prey system to become unstable with an increase in the environmental quality for the prey.

population A group of individual items. In the context of population ecology, a population is a group of individual living organisms.

population projection matrix The matrix of age-specific or stage-specific transition probabilities and natality values that is used to calculate age- or stage-specific population densities into the future.

propagule rain In the context of a metapopulation, the situation in which all subpopulations release propagules generally into the environment such that they “rain” on all subpopulations.

repeller A point or cycle that is theoretically an equilibrium but generates transients that deviate from the equilibrium position when perturbed.

reproductive value The elements of the left eigenvector of a population projection matrix.

r selection Selection for high intrinsic rate of increase, usually symbolized as r , that occurs when a population is at low density, far below carrying capacity. Typically assumed to result in higher reproductive potential and lower survivorship.

- rescue effect** In the context of a spatially subdivided population, the situation in which a subpopulation tends toward extinction but receives migrants from another subpopulation before going extinct.
- sensitivity** The degree to which the population growth rate changes as a function of changes in the elements of the projection matrix.
- separatrix** The boundary between two basins of attraction.
- stable stage (age) distribution** The proportional distribution of individuals in age or stage classes after a population has been growing (with constant transition parameters) for many generations. These proportions remain constant in perpetuity.
- strange attractor** A chaotic attractor. A region of space that attracts all trajectories but contains no attractive points or cycles.
- structural stability** A higher-level stability concept in which the qualitative nature of a system is unchanged when the parameters of the system are varied.
- structured models** Models that do not assume that all individuals in the population are identical.
- thinning laws** The relationship between the population density and population biomass in a population undergoing mortality (thinning). This relationship is most often represented as a graph of the log of the biomass per individual versus the log of the density.
- vector field** The set of vectors that determine the behavior of a dynamic system.
- yield–density relation** The yield (total biomass or the biomass of a part, e.g., seeds) of a population (especially a plant population) as a function of its density.

REFERENCES

- Alligood, K. T., T. D. Sauer, and J. Yorke. 1996. *Chaos: An introduction to dynamical systems*. New York: Springer.
- Altieri, M. A. 1987. *Agroecology: The scientific basis of alternative agriculture*. Boulder, CO: Westview Press.
- Anderson, R. 1981. Population ecology of infectious disease agents. In *Theoretical ecology: Principles and applications*, edited by R. May. Sunderland, MA: Sinauer.
- Anderson, R., and R. May. 1979. Population biology of infectious diseases: Part 1. *Nature* 280:361–367.
- . 1991. *Infectious diseases of humans: Dynamics and control*. London: Oxford Univ. Press.
- Andrewartha, H. G., and L. C. Birch. 1954. *The distribution and abundance of animals*. Chicago: Univ. of Chicago Press.
- Arditi, R., and A. A. Berryman. 1991. The biological control paradox. *Trends Ecol. Evol.* 6:32.
- Ashman, T. L. 1994. A dynamic perspective on the physiological cost of reproduction in plants. *Am. Nat.* 144:300–316.
- Bailey, N. T. J. 1975. *The mathematical theory of infectious diseases*, 2nd ed. London: Griffin.
- Baltensweiler, W. 1971. The relevance of changes in the composition of larch bud moth populations for the dynamics of its numbers. In *Dynamics of populations*, edited by P. J. den Boer and D. R. Gradwell. Wageningen: Centre for Agricultural Publishing.
- Baskerville, E., A. Dobson, T. Bedford, S. Allesina and M. Pascual. 2011. Spatial guilds in the Serengeti food web revealed by a Bayesian group model. *PloS Computational Biology* 7(12): e1002321 (doi:10.1371/journal.pcbi.1002321).
- Bascompte, J., P. Jordano, and J. M. Olesen. 2006. Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science* 312:431–433.
- Beissinger, S. R., and D. R. McCullough. 2002. *Population viability analysis*. Chicago: Univ. of Chicago Press.
- Beverton, R. J. H., and S. J. Holt. 1957. *On the dynamics of exploited fish populations*. Fishery Investigations Series 2, vol. 19. London: British Ministry of Agriculture, Fisheries, and Food.
- Bielby, J., G. M. Mace, O. R. P. Bininda-Emonds, M. Cardillo, J. L. Gittleman, K. E. Jones, C. D. L. Orme, and A. Purvis. 2007. The fast–slow continuum in mammalian life history: An empirical reevaluation. *Am. Nat.* 169(6):748–757.

- Blasius, B., A. Huppert, and L. Stone. 1999. Complex dynamics and phase synchronization in spatially extended ecological systems. *Nature* 399:354–359.
- Bleasdale, J. K. A., and J. A. Nelder. 1960. Plant population and crop yield. *Nature* 188:342.
- Boucher, D. H., and M. A. Mallona. 1997. Recovery of the rain forest tree *Vochysia ferruginea* over 5 years following Hurricane Joan in Nicaragua: A preliminary population projection matrix. *Forest Ecol. Manag.* 91:195–204.
- Boyce, M. S. 1984. Restitution of *r*- and *K*-selection as a model of density-dependent natural selection. *Annu. Rev. Ecol. Syst.* 15:427–447.
- Branch, G. M. 1975. Intraspecific competition in *Patella cochlear* Born. *J. Anim. Ecol.* 44:263–281.
- Bronstein, J. L., R. Alarcón, and M. Geber. 2006. The evolution of plant–insect mutualisms. *New Phytologist* 172:412–428.
- Bruno, J. F., J. J. Stachowicz, and M. D. Bertness. 2003. Inclusion of facilitation into ecological theory. *Trends Ecol. Evol.* 18:119–125.
- Bull, J. J. and I.-N. Wang. 2010. Optimality models in the age of experimental evolution and genomics. *J. Evol. Biol.* 23:1820–1838.
- Burns, J. H., S. P. Blomberg, E. E. Crone, J. Ehrlen, T. M. Knight, J.-P. Pichancourt, S. Ramula, G. M. Wardle, and Y. M. Buckley. 2010. Empirical tests of life-history evolution theory using phylogenetic analysis of plant demography. *J. Ecol.* 98:334–344.
- Caswell, H. 1978. A general formula for the sensitivity of population growth rate to changes in life history parameters. *Theor. Pop. Biol.* 14:215–230.
- . 2001. *Matrix population models: Construction, analysis, and interpretation*. Sunderland, MA: Sinauer.
- Caswell, H., R. Naiman, and R. Morin. 1984. Evaluation of the consequences of reproduction in complex salmonid life cycles. *Aquaculture* 43:123–143.
- Chapman, R. 1928. The quantitative analysis of environmental factors. *Ecology* 9:111–122.
- Childs, D. Z., C. J. E. Metcalf, and M. Rees. 2010. Evolutionary bet-hedging in the real world: Empirical evidence and challenges revealed by plants. *Proc. R. Soc. London B Bio.* 277:3055–3064.
- Clark, P. J., and F. C. Evans. 1954. Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology* 35:445–453.
- Cody, M. L. 1966. A general theory of clutch size. *Evolution* 20:174–184.
- Cole, L. C. 1954. The population consequences of life history phenomena. *Q. Rev. Biol.* 29:103–137.
- Costantino, R. F., R. A. Desharnais, J. M. Cushing, and B. Dennis. 1997. Chaotic dynamics in an insect population. *Science* 275:389–391.
- de Kroon, H., A. Plaisier, J. van Groenendael, and H. Caswell. 1986. Elasticity: The relative contribution of demographic parameters to population growth rate. *Ecology* 67:1427–1431.
- Deevey, E. S. Jr. 1947. Life tables for natural populations of animals. *Q. Rev. Biol.* 22:283–314.
- Dempster, J. P. 1983. The natural control of populations of butterflies and moths. *Biol. Rev.* 58:461–481.
- Dieckmann, U., R. Law, and J.A.J. Metz. 2000. *The geometry of ecological interactions: Simplifying spatial complexity*. Cambridge, UK: Cambridge Univ. Press.
- Diggle, P. J. 1983. *Statistical analysis of spatial point patterns*. New York: Academic Press.
- Doubleday, T. 1842. *The true law of population shewn to be connected with the food of the people*. London: Simpkin, Marshall.

- Ellner, S., and P. Turchin. 1995. Chaos in a noisy world: New methods and evidence from time-series analysis. *Am. Nat.* 145:343–375.
- Enquist, B. J., J. H. Brown, and G. B. West. 1998. Allometric scaling of plant energetics and population density. *Nature* 395:163–165.
- Eriksson, O., and K. Kiviniemi. 1999. Site occupancy, recruitment, and extinction thresholds in grassland plants: An experimental study. *Biol. Conserv.* 87:319–325.
- Fisher, R. A. 1937. The wave of advance of advantageous genes. *Ann. Hum. Genet.* 7:355–369.
- Flatt, T., and A. Heyland. 2012. *Mechanisms of life history evolution: The genetics and physiology of life history traits and trade-offs*. Oxford, UK: Oxford Univ. Press.
- Fortin, M. J., and M. R. T. Dale. 2005. *Spatial analysis: A guide for ecologists*. New York: Wiley.
- Franco, M., and J. Silvertown. 1996. Life history variation in plants: An exploration of the fast-slow continuum hypothesis. *Phil. Trans. R. Soc. London B* 351:1341–1348.
- . 2004. A comparative demography of plants based upon elasticities of vital rates. *Ecology* 85:531–538.
- Gause, G. F. 1934. *The struggle for existence*. Baltimore: Williams and Wilkins.
- Getz, W. M., and R. G. Haight. 1989. Population harvesting: *Demographic models of fish, forest, and animal resources*. Monographs in Population Biology 27. Princeton, NJ: Princeton Univ. Press.
- Golet, G. H., J. A. Schmutz, D. B. Irons, et al. 2004. Determinants of reproductive costs in the long-lived Black-legged Kittiwake: A multiyear experiment. *Ecol. Monogr.* 74(2):353–372.
- Gotelli, N. J. 1991. Metapopulation models: The rescue effect, the propagule rain, and the core-satellite hypothesis. *Am. Nat.* 138:768–776.
- Guckenheimer, J., G. Oster, and A. Ipaktchi. 1977. The dynamics of density dependent population models. *J. Math. Biol.* 4:101–147.
- Hairton, N. G., D. W. Tinkle, and H. M. Wilbur. 1970. Natural selection and the parameters of population growth. *J. Wildlife Manage.* 34:681–690.
- Hanski, I. 1982. Dynamics of regional distributions: The core and satellite species hypothesis. *Oikos* 38:210–221.
- . 1999. *Metapopulation biology: Ecology, genetics, and evolution*. San Diego: Academic Press.
- Hanski, I., and O. G. Gaggiotti, eds. 2004. *Ecology, genetics, and evolution of metapopulations*. Burlington, MA: Elsevier Academic Press.
- Hanski, I., and D. Simberloff. 1997. The metapopulation approach, its history, conceptual domain, and application to conservation. In *Metapopulation biology: Ecology, genetics, and evolution*, edited by I. Hanski and M. Gilpin. San Diego: Academic Press.
- Harrison, S. 1991. Local extinction in a metapopulation context. In *Metapopulation dynamics: Empirical and theoretical investigations*, edited by M. Gilpin and I. Hanski. London: Academic Press.
- Harrison, S., and A. D. Taylor. 1997. Empirical evidence for metapopulation dynamics. In *Metapopulation biology: Ecology, genetics, and evolution*, edited by I. Hanski and M. Gilpin. San Diego: Academic Press.
- Hassell, M. P. 1975. Density-dependence in single-species populations. *J. Anim. Ecol.* 45:283–296.
- . 1978. *The dynamics of arthropod predator-prey systems*. Princeton Monographs in Population Biology. Princeton, NJ: Princeton Univ. Press.
- Hastings, A., and T. Powell. 1991. Chaos in a three species food chain. *Ecology* 72:896–903.

- Hastings, A., C. L. Hom, S. Ellner, P. Turchin, and H. C. J. Godfray. 1993. Chaos in ecology: Is mother nature a strange attractor? *Annu. Rev. Ecol. Syst.* 24:1–33.
- Hawkins, B. A., and H. V. Cornell. 1999. *Theoretical approaches to biological control*. Cambridge, UK: Cambridge Univ. Press.
- Hiroi, T., and M. Monsi. 1966. Dry-matter economy of *Helianthus annuus* communities grown at varying densities and light intensities. *J. Fac. Sci. U. Tokyo* 9:241–285.
- Hoeksema, J., V. Bala Chaudhary, C. A. Gehring, N. C. Johnson, J. Karst, R. T. Koide, A. Pringle, C. Zabinski, J. D. Bever, J. C. Moore, G. W. T. Wilson, J. N. Klironomos, and J. Umbanhowar. 2010. A meta-analysis of context-dependency in plant response to inoculation with mycorrhizal fungi. *Ecol. Lett.* 13:394–407.
- Holland, J. N., and D. L. DeAngelis. 2010. A consumer–resource approach to the density-dependent population dynamics of mutualism. *Ecology* 91:1286–1295.
- Holland, J. N., D. L. DeAngelis, and J. L. Bronstein. 2002. Population dynamics and mutualism: Functional responses of benefits and costs. *Am. Nat.* 159:231–244.
- Holling, C. S. 1959. Some characteristics of simple types of predation and parasitism. *Can. Entomol.* 91:385–398.
- Howard, L. O., and W. F. Fiske. 1911. *The importation into the United States of the parasites of the Gipsy moth and the brown-tail moth: A report of progress, with some consideration of previous and concurrent efforts of this kind*. U.S. Department of Agriculture, Bureau of Entomology, Bulletin 91:1–344.
- Hubbell, S. P. 2001. *The unified neutral theory of biodiversity and biogeography*. Monographs in Population Biology. Princeton NJ: Princeton Univ. Press.
- Huffaker, C. B. 1958. Experimental studies on predation: Dispersion factors and predator–prey oscillations. *Hilgardia* 27:343–383.
- Jackson, E. A. 1991. *Perspectives of nonlinear dynamics*, vols. 1 and 2. Cambridge, UK: Cambridge Univ. Press.
- Jackson, L. L., and C. L. Dewald. 1994. Predicting evolutionary consequences of greater reproductive effort in *Tripsacum dactyloides*, a perennial grass. *Ecology* 75:627–641.
- Jackson, W. 1980. *New roots for agriculture*. San Francisco: Friends of the Earth.
- Kays, S., and J. L. Harper. 1974. The regulation of plant and tiller density in a grass sward. *J. Ecol.* 62:97–10.
- Keeling, M. J., and P. Rohani. 2008. *Modeling infectious diseases in humans and animals*. Princeton, NJ: Princeton Univ. Press.
- Kermack, W. O., and A. G. McKendrick. 1927. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. London A-Cont.* 115:700–721.
- Lack, D. 1947. The significance of clutch-size, parts 1 and 2. *Ibis* 89:302–352.
- Lande, R. 1982. A quantitative genetic theory of life history evolution. *Ecology* 63:607–615.
- Lefkovich, L. P. 1965. The study of population growth in organisms grouped by stages. *Biometrics* 21:1–18.
- Leslie, P. H. 1945. On the use of matrices in certain population mathematics. *Biometrika* 33:183–212.
- Levins, R. 1968. *Evolution in changing environments*. Princeton, NJ: Princeton Univ. Press.
- . 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bull. Entomol. Soc. Am.* 15:237–240.
- Levins, R., and J. H. Vandermeer. 1989. The agroecosystem embedded in a complex ecological community. In *Agroecology*, edited by C. R. Carroll, J. H. Vandermeer, and P. Rosset, 341–362. New York: McGraw-Hill.
- Lewis, E. G. 1942. On the generation and growth of a population. *Sankhyaa* 6:93–96.

- Lewontin, R. C. 1969. The meaning of stability. *Brookhaven Symp. Biol.* 22:13–24.
- Łomnicki, A. 1988. *Population ecology of Individuals*. Princeton, NJ: Princeton Univ. Press.
- . 2009. Scramble and contest competition, unequal resource allocation, and resource monopolization as determinants of population dynamics. *Evol. Ecol. Res.* 11:371–380.
- Lotka, A. J. 1926. *Elements of physical biology*. Baltimore: Williams and Wilkins.
- MacArthur, R. 1970. Species packing and competitive equilibria for many species. *Theor. Pop. Biol.* 1:1–11.
- . 1972. *Geographical ecology*. New York: Harper and Row.
- MacArthur, R. H., and E. O. Wilson. 1967. *The theory of island biogeography*. Princeton, NJ: Princeton Univ. Press.
- Macdonald, G. 1957. *The epidemiology and control of malaria*. London: Oxford Univ. Press.
- Malthus, T. 1960 [1830]. *A summary view of the principle of population*. New York: New American Library.
- Marmontel, M., S. R. Humphrey, and T. J. O'Shea. 1997. Population viability analysis of the Florida Manatee (*Trichechus manatus latirostris*), 1976–1991. *Conserv. Biol.* 11:467–481.
- May, R. M. 1973. *Stability and complexity in model ecosystems*. Princeton Monographs in Population Biology. Princeton, NJ: Princeton Univ. Press.
- . 1974. Biological populations with nonoverlapping generations: Stable points, stable cycles, and chaos. *Science* 186:645–647.
- . 1977. Thresholds and breakpoints in ecosystems with a multiplicity of stable states. *Nature* 269:471–477.
- . 1981. *Theoretical ecology: Principles and applications*, 2nd ed. Sunderland, MA: Sinauer.
- McCann, K., and P. Yodzis. 1994. Nonlinear dynamics and population disappearances. *Am. Nat.* 144:873–879.
- Méndez-Alonso, R., H. Hernández-Trejo, and J. López-Portillo. 2012. Salinity constrains size inequality and allometry in two contrasting mangrove habitats in the Gulf of Mexico. *J. Trop. Ecol.* 28(2):171–179.
- Metcalf, C. J. E., and S. Pavard. 2006. Why evolutionary biologists should be demographers. *Trends Ecol. Evol.* 22:205–212.
- Milne, A. 1961. Definition of competition among animals. In *Mechanisms in biological competition*, edited by F. L. Milthorpe. *Symp. Soc. Exp. Biol.* 15:40–61. New York: Academic Press.
- Morales, H. 1998. Indigenous methods of pest control in the Guatemalan highlands. Ph.D. diss., Univ. of Michigan, Ann Arbor.
- Murie, O. 1944. *The wolves of Mt. McKinley*. Fauna Ser. 5. Washington, DC: U.S. Department of the Interior, National Park Service.
- Murdoch, W. W., S. O. Swarbrick, R. F. Luck, S. Walde, and D. S. Yu. 1996. Refuge dynamics and meta population dynamics: An experimental test. *Am. Nat.* 147:424–444.
- Nicholson, A. J. 1933. The balance of animal populations. *J. Anim. Ecol.* 2:123–178.
- . 1954. An outline of the dynamics of animal populations. *Aust. J. Zool.* 2:9–65.
- Nicholson, A. J., and V. A. Bailey 1935. The balance of animal populations, Part I. *Proc. Zool. Soc. London* 105:551–598.
- Noy-Meir, E. 1975. Stability of grazing systems: An application of predator–prey graphs. *J. Ecol.* 63:459–481.
- Oli, M. K., and F. S. Dobson. 2003. The relative importance of life-history variables to population growth rate in mammals: Cole's prediction revisited. *Am. Nat.* 161:422–440.

- Paige, K. N., and T. G. Whitham. 1987. Overcompensation in response to mammalian herbivory: The advantage of being eaten. *Am. Nat.* 129:407–416.
- Palmer, M. W. 1994. Variation in species richness: Towards a unification of hypotheses. *Folia Geobot. Phytotx.* 29:511–530.
- Pascual, M. and J. Dunne (editors). 2005. *Ecological Networks: Linking Structure to Dynamics in Food Webs*. (Proceedings of workshop at the Santa Fe Institute). SFI and Oxford Univ. Press.
- Pascual, M., and S. A. Levin. 1999. From individuals to population densities: Searching for the intermediate scale of nontrivial determinism. *Ecology* 80:2225–2236.
- Pascual, M., P. Mazzega, and S. Levin. 2001. Oscillatory dynamics and spatial scale: The role of noise and unresolved pattern. *Ecology* 82:2357–2369.
- Pearl, R., and L. J. Reed. 1920. On the rate of growth of the population of the United States since 1790 and its mathematical representation. *Proc. Natl. Acad. Sci. USA* 6:275–288.
- Perfecto, I., and J. Vandermeer. 2010. The agroecological matrix as alternative to the land-sparing/agriculture intensification model. *Proc. Natl. Acad. Sci. USA* 107:5786–5791.
- Pervez, A., and Omkar. 2005. Functional responses of coccinellid predators: An illustration of a logistic approach. *J. Insect Sci.* 5:1–5.
- Pianka, E. 1970. On *r* and *K* selection. *Am. Nat.* 104:592–597.
- Ramula, S., T. M. Knight, J. H. Burns, and Y. M. Buckley. 2008. General guidelines for invasive plant management based on comparative demography of invasive and native plant populations. *J. Appl. Ecol.* 45(4):1124–1133.
- Ricker, W. E. 1954. Stock and recruitment. *J. Fish. Res. Board Can.* 11:559–623.
- Roff, D. A. 1992. *The evolution of life histories: Theory and analysis*. New York: Chapman and Hall.
- Rose, M. R., and B. Charlesworth. 1981. Genetics of life history in *Drosophila melanogaster*, part 2: Exploratory selection experiments. *Genetics* 97:187–196.
- Rosenzweig, M. L., and R. H. MacArthur. 1963. Graphic representation and stability conditions of predator–prey interactions. *Am. Nat.* 97:209–223.
- Rosenzweig, M. P. 1971. Paradox of enrichment: Destabilization of exploitation ecosystems in ecological time. *Science* 171:385–387.
- Ross, R. 1911. *The prevention of malaria*, 2nd ed. London: Murray.
- Schaffer, W. M. 1974. Optimal reproductive effort in fluctuating environments. *Am. Nat.* 108:783–790.
- . 1985. Order and chaos in ecological systems. *Ecology* 66:93–106.
- Schaffer, W. M., and M. D. Gadgil. 1975. Selection for optimal life histories in plants. In *Ecology and evolution of communities*, edited by M. L. Cody and J. M. Diamond. Cambridge, MA: Harvard Univ. Press.
- Schaffer, W. M., and M. V. Schaffer. 1977. The adaptive significance of variations in reproductive habit in the Agavaceae. In *Evolutionary ecology*, edited by B. Stonehouse and C. M. Perrins. London: Macmillan.
- Scheffer, M. 2009. *Critical transitions in nature and society*. Princeton, NJ: Princeton Univ. Press.
- Scheffer, M., J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Vanes, M. Rietkerk, and G. Sugihara. 2009. Early-warning signals for critical transitions. *Nature* 461:53–59.
- Scheffer, M., S. Carpenter, T. M. Lenton, J. Bascompte, W. Brock, V. Dakos, J. van den Koppel, I. van de Leemput, S. Levin, E. H. van Nes, M. Pascual, and J. Vandermeer. 2012. Anticipating critical transitions. *Science* 338:344–348.
- Shinozaki, K., and T. Kira. 1956. Intraspecific competition among higher plants, VII: Logistic theory of the C-D effect. *J. Inst. Polytech. Osaka City Univ.*, ser. D 7:35–72.

- Silvertown, J. 1980. The evolutionary ecology of mast seeding in trees. *Biol. J. Linn. Soc.* 14:235–250.
- Silvertown, J., M. Franco, and J. L. Harper, eds. 1997. *Plant life histories: Ecology, phylogeny, and evolution*. Cambridge, UK: Cambridge Univ. Press.
- Smith, H. S. 1935. The role of biotic factors in the determination of population densities. *J. Econ. Entomol.* 28:873–898.
- Solé, R., and B. Goodwin. 2000. *Signs of life: How complexity pervades biology*. New York: Basic Books.
- Soule, J., D. Carré, and W. Jackson 1990. Ecological impact of modern agriculture. In *Agroecology*, edited by C. R. Carroll, J. H. Vandermeer, and P. Rosset, 165–188. New York: McGraw-Hill.
- Stearns, S. 1992. *The evolution of life histories*. Oxford, UK: Oxford Univ. Press.
- Stiven, A. E. 1967. The influence of host population space in experimental epizootics caused by *Hydramoeba hydroxena*. *J. Invert. Pathol.* 9:536–545.
- Strogatz, S. H. 2001. *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*. Studies in Nonlinearity. Cambridge, UK: Westview Press.
- Strong, D. 1986. Density-vague population change. *Trends Ecol. Evol.* 1:39–42.
- Thompson, W. R. 1928. A contribution to the study of biological control and parasite introduction in continental areas. *Parasitology* 20:90–112.
- Tilman, D. 1982. *Resource competition and community structure*. Princeton, NJ: Princeton Univ. Press.
- . 1997. Community invasibility, recruitment limitation, and grassland biodiversity. *Ecology* 78:81–92.
- Turing, A. M. 1952. The chemical basis of morphogenesis. *Phil. Trans. R. Soc. London B* 237:37–72.
- Turnbull, L. A., C. D. Philipson, D. W. Purves, R. L. Atkinson, J. Cunniff, A. Goode-nough, Y. Hautier, J. Houghton, T. R. Marthews, C. P. Osborne, C. Paul-Victor, K. E. Rose, P. Saner, S. H. Taylor, F. I. Woodward, A. Hector, and M. Rees. 2012. Plant growth rates and seed size: A re-evaluation. *Ecology* 93(6):1283–1289.
- Turner, M. D., and D. Rabinowitz. 1983. Factors affecting frequency distributions of plant mass: The absence of dominance and suppression in *Festuca paradoxa*. *Ecology* 64:469–475.
- United Nations, Department of Economic and Social Affairs. 2011. *World Population Prospects, the 2010 Revision*. Population Division, Population Estimates and Projections Section. <http://esa.un.org/unpd/wpp/Excel-Data/population.htm>.
- Utida, S. 1957. Population fluctuation: An experimental and theoretical approach. *Cold Spring Harbor Symp. Quant. Biol.* 22:139–151.
- Vandermeer, J. H. 1969. The competitive structure of communities: An experimental approach using protozoa. *Ecology* 50:362–371
- . 1975. On the construction of the population projection matrix for a population grouped in unequal stages. *Biometrics* 31:239–242.
- . 1989. *The ecology of intercropping*. Cambridge, UK: Cambridge Univ. Press.
- Vandermeer, J. H., and D. Boucher. 1978. Varieties of mutualistic interaction in population models. *J. Theor. Biol.* 74:549–558.
- Vandermeer, J. H., and P. Yodzis. 1999. Basin boundary collision as a model of discontinuous change in ecosystems. *Ecology* 80:1817–1827.
- Vandermeer, J., I. Perfecto, and S. Philpott. 2010. Ecological complexity and pest control in organic coffee production: Uncovering an autonomous ecosystem service. *BioScience* 60:527–537.
- Vannette, R., and M. D. Hunter. 2011. Plant defence theory re-examined: Nonlinear expectations based on the costs and benefits of resource mutualisms. *J. Ecol.* 99:66–76.

- Verhulst, P. F. 1838. Notice sur la loi que la population suit dans son accroissement. *Corresp. Math. Phys.* 10:113–121.
- Volterra, V. 1926. Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Memorie della accademia Lincei* 2:31–113.
- von Forester, H., P. M. Mora, and L. W. Amiot. 1960. Doomsday: Friday, 13 November, A.D. 2026. *Science* 132:1291–1295.
- Warne, R. W., and E. L. Charnov. 2008. Reproductive allometry and the size–number trade-off for lizards. *Am. Nat.* 172(3):E80–E98.
- Warren, K. S., and V. A. Newill. 1967. Schistosomiasis—A bibliography of the world's literature from 1852 to 1962, vols. 1 and 2. Cleveland, OH: Western Reserve Univ. Press.
- Watkinson, A. R. 1980. Density-dependence in single-species populations of plants. *J. Theor. Biol.* 83:345–357.
- Weiner, J. 1986. How competition for light and nutrients affects size variability in *Ipomoea tricolor* populations. *Ecology* 67:1425–1427.
- . 1990. Asymmetric competition in plant populations. *Trends Ecol. Evol.* 5:360–364.
- Weiner, J., and S. C. Thomas, 1986. Size variability and competition in plant monocultures. *Oikos* 47:211–222.
- West, G. B., J. H. Brown, and B. J. Enquist. 1997. A general model for the origin of allometric scaling laws in biology. *Science* 276:122.
- Westoby, M. 1984. The self-thinning rule. *Adv. Ecol. Res.* 14:167–225.
- Westoby, M., M. Leishman, and J. Lord. 1996. Comparative ecology of seed size and dispersal. *Phil. Trans. R. Soc. London B* 351:1309–1318.
- Wilbur, H. M., D. W. Tinkle, and J. P. Collins. 1974. Environmental certainty, trophic level, and resource availability in life history evolution. *Am. Nat.* 108:805–817.
- Wiley, R. W., and Heath, S. B. 1969. The quantitative relationships between plant population and crop yield. *Adv. Agron.* 21:281–321.
- Yoda, K., T. Kira, H. Ogawa, and H. Hozumi, H. 1963. Self-thinning in overcrowded pure stands under cultivated and natural conditions. *J. Inst. Polytech. Osaka City Univ.*, ser. D 14:107–129.
- Yodzis, P. 1978. *Competition for space and the structure of ecological communities*. Berlin: Springer.

INDEX

Page numbers for entries occurring in figures are followed by an *f* and those for entries in tables, by a *t*.

- acacia trees, 126, 137
- action threshold, 2
- activators, in spatial patterns, 142, 143*f*
- adults: in age-structured projection matrices, 30; in stage-structured projection matrices, 40–41
- age distribution, stable, 34–37, 35*f*, 45–47
- age distribution vector, 31–32
- age-structured population(s), static life tables of, 73–74
- age-structured population projection matrices, 30–39; density dependence in, 31–32, 36, 48–50, 50*f*; stable age distribution in, 34–37, 35*f*, 45–47
- agriculture: intercropping in, 211–12, 213*f*; natural systems, 67
- AIDS, 242
- allometric effect, 21
- American basswood, 127, 150–51
- American beech, 127, 150–51
- Anderson, R., 194
- Andrewartha, H. G., 10
- animals: contest and scramble competition in, 53, 54; optimal reproductive schedules for, 69; thinning laws in, 24, 27. *See also specific types*
- annual plants, 67, 72
- ants, 208–9
- aphids, 6–8, 7*f*, 7*t*, 8*f*, 16
- Armillaria gallica*, 50
- artificial selection, 66
- Ashman, T. L., 66
- attractors: in bifurcation diagrams, 116–22; definition of, 84, 84*f*; eigenvalues and, 92–97, 96*f*; in interspecific competition, 206, 206*f*; in logistic map, 106–10; in one-dimensional maps, 97–106; periodic, 87–89, 88*f*, 90; physical models of, 85–90, 85*f*–89*f*; point (See point attractors); in predator–prey interactions, 172–73, 173*f*; strange (See strange attractors); structural stability and, 110–15, 111*f*; types of, 85–91; use of term, 84–85
- autocatalytic species, 142
- autonomous pattern formation, 137, 141
- Azuki bean weevil (*Callosobruchus chinensis*), 157*f*, 158
- Bailey, V. A., 9–10, 176, 179–80, 182, 186
- balance. *See* equilibrium
- basic reproductive rate of disease, 192–93, 194, 197
- basin boundary collision, 115, 116*f*, 124
- basins of attraction: definition of, 90; edges of, 90; in logistic map, 108–10, 109*f*
- beaker and magnet model of dynamics, 85–86, 86*f*, 89, 90
- beetles, 56–57, 162, 162*f*, 213–14
- beets, 18*f*
- bet hedging, 71
- Beverton, R.J.H., 29
- bifurcation, 112–22; diagrams of, 116–22, 117*f*–20*f*; Hopf, 113; period-doubling, 113, 114, 118–19, 118*f*; in predator–prey interactions, 156; saddle-node, 113–14, 113*f*, 115*f*; types of, 113–15

- bifurcation points, 112–14, 113f, 124, 156
- biodiversity, loss of, 241–42
- biological control, paradox of, in predator–prey interactions, 170, 170f, 176
- biomass: in thinning process, 22–27, 24f, 25f, 26f, 27f; in yield–density relationship, 19–21
- Birch, L. C., 10
- birds, spatial distribution of, 128–30, 128f–30f
- birth rate: density dependence and, 17–22; in exponential growth, 6; in life history analysis, 65, 70–71; in predator–prey interactions, 154, 164–65; resource availability and, 217–18, 217f, 218f; in static life tables, 73
- blackbirds, 128, 128f
- Bleasdale, J.K.A., 21, 22, 28–29
- blue-sky bifurcation. *See* saddle-node bifurcation
- Boucher, D. H., 74–75
- Branch, G. M., 53
- break point, 172
- Bull, J. J., 68
- butterflies, 146–47
- Callosobruchus chinensis* (Azuki bean weevil), 157f, 158
- cannibalism, 56, 120
- carrying capacity: calculation of, 16; definition of, 16; density and, 16–17, 28; in interspecific competition, 199–202, 202f, 207–8; vs. intrinsic rate of natural increase, selection for, 63, 64–65; in life history analysis, 63–65; in metapopulations, 145–46; in mutualism, 226, 228–30, 232, 233f, 237; origins of concept, 9; paradox of enrichment and, 168–69, 169f; in predator–prey interactions, 168–69, 171, 173–74
- Caswell, H., 47
- catastrophic factors, 9
- centers, 158. *See also* neutral stability
- chaos: period-doubling route to, 119; sensitive dependence on initial conditions in, 124–25; use of term, 124
- chaotic attractors. *See* strange attractors
- chaotic behavior, of strange attractors, 89–91, 106, 125
- chaotic populations: density of, 10–11; vs. random populations, 10
- Chapman, R., 9
- Charlesworth, B., 66, 67f
- chess metaphor, 207
- Chlorohydra viridissima*, 188, 189f
- Clark, P. J., 135
- clumping, 126–29, 133–34, 143
- clustering, 134–37
- Coccinella transversalis*, 162, 162f
- Cody, M. L., 64
- coexistence, competitive, 206–11, 208f
- cohort life tables. *See* dynamic life tables
- Cole, L. C., 62, 64
- colonization, of patches, 145
- column vectors, 45, 57–58
- community ecology: competitive exclusion principle in, 211; species diversity in, 240; stability and equilibrium in, 81
- competition, interspecific, 198–224; classical analysis of outcomes of, 204–7, 206f; competitive exclusion principle in, 207–11, 208f; competitive production principle in, 211–12; definition of, 199; first principles of, 198–203; graphical approach to, 216–24; history of concept, 198; Lotka–Volterra equations for, 199–209, 212–14; mechanistic approach to, 213–16; niches in, 209–11; phenomenological approach to, 212–14; significance of, 152. *See also specific types*
- competition, intraspecific, 9–29; definition of, 9; as density dependence, 9, 17; development of concept, 9–11; logistic equation and, 13–17; in optimal reproductive schedules, 70; in size distributions, 51–55, 53f, 54f; types of, 53–54; yield–density relationship and, 17–22
- competition coefficients, 20–21, 201–2, 202f, 209–10
- competitive coexistence, 206–11, 208f
- competitive exclusion principle, 207–11, 208f
- competitive production principle, 211–12
- complexity, of systems, stability related to, 81–82, 122
- connectivity, of systems, 81–82
- constant final yield, law of, 19
- constant predator model, 109, 113–14, 115
- consumer–resource interactions, definition of, 152. *See also* competition; predator–prey interactions
- consumption vectors, 220–22, 220f, 221f, 223f
- contest competition, 53, 54
- continuous time, logistic equation in, 11
- core–satellite hypothesis, 148
- corn, 19
- Costantino, R. F., 56, 120, 120f, 121f
- Cramer's rule, 61
- Crematogaster carinata*, 208–9
- crude attack rate, 180–81

- Dall mountain sheep (*Ovis dalli*), 73–74, 74t
Daphnia pulicaria, 187–88, 189
 Darwin, Charles, 198
 DeAngelis, D. L., 235
 death rate: and density dependence, 22–27; in exponential growth, 6; in life history analysis, 65; in predator–prey interactions, 154; resource availability and, 217–18, 217f, 218f; in size distributions, 54–55; in static life tables, 73–74
 demography, projection matrices in analysis of, 78–80
 Dempster, J. P., 10
 density, debate over dependent vs. independent factors in, 9–11
 density dependence, 9–29; in age-structured populations, 31–32, 36, 48–50, 50f; and birth rate, 17–22; and carrying capacity, 16–17, 28; and death rate, 22–27; debate over, 9–11; in discrete time models, 11, 28–29; as intraspecific competition, 9, 17; logistic equation and, 11, 13–17; in mutualism, 229–30; in predator–prey interactions, 154, 158–68, 160f, 167f, 171; in r–K selection, 64; in size-distributed populations, 50–55, 53f; in structured population models, 36, 48–57; yield–density relationship and, 11, 17–22, 19f
 density independence, 2–11
 density vagueness, 10
 detached coefficient matrix, 57, 60–61
 determinants, of matrices, 59–61
 Dewald, C. L., 67
 diffusion, in formation of spatial patterns, 137–42, 139f
 direct disease transmission, 188–94
 discrete rate of population growth. *See* finite rate of population growth
 discrete time: density dependence in, 11, 28–29; predator–prey interactions in, 176–86, 179f, 182–85f
 disease ecology, 187–97; direct transmission in, 188–94; indirect transmission in, 194–97; role of population ecology in study of, 240, 242
 diversity: bio-, loss of, 241–42; species, 240
 dominant eigenvalues, 37–39, 94, 95
Drosophila melanogaster, 66, 67f
 dynamic boundedness, 106
 dynamic (cohort) life tables, 73, 74–76
 dynamics, population, 81–125; eigenvalues in, 92–97; physical models of, 85–90, 85f–89f, 94–95, 95f; problems with traditional theories on, 81–82, 122–23. *See also* equilibrium; stability
 Ebola virus, 242
 ecological niches, 209–11
 ecology: fundamental principles of, xviii; origins and development of field, xvii–xviii; paradoxes of, 122
 ecosystem management: paradox of biological control in, 170; paradox of enrichment in, 168–70; role of population ecology in, 242; stability and equilibrium in, 81; stage-structured projection matrices in, 42–43
 eigenvalues, 45–48; definition of, 46, 92, 92f; dominant, 37–39, 94, 95; in dynamic analysis, 92–97, 104, 105f; of projection matrices, 45–48, 94
 eigenvectors, 45–48
 elasticity, of projection matrices, 47–48; applications of, 75–76, 78–80
 encounter time, 163–64
 endemic diseases, 190, 192, 194
 energy allocation, to reproduction, 68–72, 69f
 Enquist, B. J., 27
 enrichment, paradox of, in predator–prey interactions, 168–70, 169f
 environmental resistance, 9
 environmental space, in interspecific competition, 201–2, 202f
Eotetranychus sexmaculatus (six-spotted mites), 157f, 158
 epidemic diseases, 190–91, 192–93
 epidemiology, 187–97, 240; definition of, 187; of direct transmission, 188–94; of indirect transmission, 194–97; role of population ecology in, 240, 242; SIR model of, 190–94, 190f, 192f, 195
 equations, uses of, 3. *See also* specific equations
 equilibrium (balance), 81–125; in bifurcation diagrams, 116–22; centrality of concept, 81; definition of, 83; in density dependence, 10; in disease ecology, 190, 196; eigenvalues and, 92–97; in interspecific competition, 205–8, 216–17; intuitive ideas of, 82, 83–91; in logistic map, 106–10; in metapopulations, 146; in one-dimensional maps, 97–106; in predator–prey interactions, 156; problems with traditional theories on, 81–82, 122–23; structural stability and, 110–15

- equilibrium points: definition of, 15, 83; stable, 83–84, 84f (*See also* attractors); unstable, 83–84, 84f (*See also* repellers)
- essential resources, 219–20, 219f, 220f
- Euler's constant, 4, 186
- Evans, F. C., 135
- evolution: interspecific competition in, 198; of mutualisms, 225. *See also* natural selection
- exclusion, competitive, 207–11, 208f
- exercises, how to use, xviii–xix
- exponential equation, 2–8, 5f, 28; differential form of, 5; and eigenvalues, 92–93; integrated form of, 6; as one-dimensional graph, 99–102, 101f
- exponential growth, 2–9; definition of, 9; equation for (*See* exponential equation); logistic equation and, 13–14; with stable age distribution, 35–36, 35f
- extinction: in disease ecology, 190; in logistic map, 107–8, 107f; in metapopulations, 143, 144–49, 149f; in population viability analysis, 76–78
- facilitation, 225–38; definition of, 225; in optimal reproductive schedules, 70
- facultative factors, 9
- facultative mutualism, 228, 228f, 231–33, 232f, 234f
- fecundity: in age-structured projection matrices, 31–32, 49; in life history analysis, 63, 66, 69–70; in stage-structured projection matrices, 40
- finite (discrete) rate of population growth, 3–4, 35, 36
- Fisher, R. A., 140
- Fiske, W. F., 9
- fixed points, 84. *See also* equilibrium points
- Florida manatee (*Trichechus manatus latirostris*), 76–78, 77t
- flour beetles, 56–57, 213–14
- functional response: in predator–prey interactions, 161–68, 167f, 171–76, 171f, 173f, 177f; type I, 163f, 175, 176f; type II, 162–63, 162f, 163f, 165f, 175, 176f; type III, 175–76, 176f
- fungi, mycorrhizal, 235–38, 237f
- Gadgil, M. D., 68
- Gause, G. F., xvii, 9, 198
- genetics, in life history evolution, 66, 68
- geometry, in three-halves thinning law, 25–27
- Gini coefficient, 52, 53
- Gotelli, N. J., 149
- grasses, life history evolution in, 67
- growth, population: density-dependent (*See* density dependence); density-independent, 2–11; exponential (*See* exponential growth); finite (discrete) rate of, 3–4, 35, 36
- habitat: clustering of, 137; of metapopulations, 142–49
- Haeckel, Ernst, xvii
- handling time, 163–64
- Hanski, I., 146, 148
- hares, 156–57, 157f
- Harrison, S., 143, 147
- Hassell, M. P., 29
- Helianthus annuus*, 24f
- herbivory, as size dependent, 51
- herd immunity, 193, 194
- Heterospilus*, 157f, 158
- hill model of dynamics, 87–88, 88f, 89, 90
- Holland, J. N., 235
- Holling, C. S., 162, 175
- Holling disk equation, 162–64
- Holt, S. J., 29
- Hopf bifurcation, 113
- hosts. *See* disease ecology
- Howard, L. O., 9
- Hudson Bay Company, 156–57, 157f
- Huffaker, C. B., 158
- human population, 1, 2, 9, 12
- Hunter, M. D., 235
- Hydramoeba hydroxena*, 188, 189f
- hyperspace, 123, 123f
- hysteresis, 114
- identity matrix, 38, 59
- immunity, 190, 193, 194
- indirect disease transmission, 194–97
- individuals, classes of, 30
- insects, stage-structured projection matrices for, 40. *See also specific types*
- intercropping, 211–12, 213f
- interspecific interactions: significance of, 153–54; types of, 153. *See also specific types*
- intraspecific competition. *See* competition, intraspecific
- intrinsic rate of natural increase: vs. carrying capacity, selection for, 63, 64–65; in exponential equation, 6–7, 14; in logistic equation, 16, 22; in projection matrices, 36–38

- invariant loops, 183–84, 184f, 185f, 186
- invasive plants, demography of, 78–80, 79f
- Ipomoea tricolor*, 54f
- Ipomopsis aggregata* (scarlet gilia), 70
- island biogeography theory, 144–45
- islands, spatial patterns on, 143
- isoclines: definition of, 155; in disease ecology, 195–96, 196f; in interspecific competition, 203–9, 204f–6f; in mutualisms, 226, 227f; in predator–prey interactions, 155–56, 155f, 156f, 159, 160f, 165, 165f, 166–67, 168f, 172–76, 173f–75f
- i-stage models, 44. *See also* stage-structured population
- iteroparity, 69–72, 69f
- Jackson, L. L., 67
- Jackson, W., 67
- Kermack, W. O., 190–91
- Kira, T., 18–19, 19f, 21, 22
- K selection, 63, 64–65, 65t
- Lack, D., 62, 64
- ladybird beetles, 162, 162f
- land equivalent ratio (LER), 211–12, 213f
- larch bud moth, 170
- Lefkovitch, L. P., 40
- Lefkovitch matrices, 40
- LER. *See* land equivalent ratio
- Leslie (Leslie–Lewis) matrix, 31, 39–40. *See also* projection matrices
- Levins, R., xvii, 143, 144–45, 148, 149, 190
- Lewontin, R. C., 106
- life history evolution, 62–72; classic studies in, 62, 64; comparative analyses of, 78; cost of reproduction in, 63, 65, 66–67; optimal reproduction schedules in, 67–72; *r*–*K* selection in, 63, 64–65, 65t; trade-offs in, 63, 65
- life history traits: definition of, 63; examples of, 63
- life tables: dynamic (cohort), 73, 74–76; static, 73–74
- lily pads, 2–3
- limit cycles, 90. *See also* periodic attractors
- linear equations: Cramer's rule for, 61; in matrix manipulations, 57, 60–61
- logistic difference equation, 28
- logistic equation, 13–17; continuous vs. discrete form of, 11; in disease ecology, 188–91; and eigenvalues, 93–94, 93f; in interspecific competition, 199, 201–2; origins of, 9; in predator–prey interactions, 158–59
- logistic map, 28–29; and age-structured population projection matrices, 33; basins of attraction in, 108–10; bifurcation diagram for, 116–18, 118f, 119f; definition of, 28; stability and equilibrium in, 106–10, 112, 112f, 113f
- Lolium perenne*, 26f
- loops, invariant, 183–84, 184f, 185f, 186
- Lotka, A. J., 153, 155, 176, 198, 211, 214
- Lotka–Volterra equations: competitive exclusion principle and, 211; density dependence and, 159, 162, 166; discrete form of, 176–79, 179f; for interspecific competition, 199–209, 212–14; for mutualisms, 225–26; origins of, xvii; oscillation and, 158; for predator–prey interactions, 154, 158–66, 176–79
- lynx, 156–57, 157f
- MacArthur, Robert, xvii, 64, 171, 173–74, 174f, 214, 216
- Macdonald, G., 194
- mainland–island populations, 143, 144f, 147
- maize, 18f, 21
- malaria, 194–97, 242
- Mallona, M. A., 74–75
- Malthus, Thomas, 9
- manatees, 76–78, 77t
- mangroves, 51f
- maps. *See* logistic map; one-dimensional maps
- marble models of dynamics, 85, 85f, 87–88, 88f, 94–95
- Marmontel, M., 76–78, 77t
- mass action assumption, 188, 191
- mathematics, role in ecology, xvii
- matrices: addition of, 58–59; definition of, 57; determinant of, 59–61; multiplication of, 57–58; square, 37–38, 42–43, 46, 57–60; subtraction of, 58–59. *See also* projection matrices
- maximum sustained yield, 17
- May, Robert, xvii, 82, 194
- McKendrick, A. G., 190–91
- metabolic scaling relationships, 27
- metapopulations: assumptions about, 146–48; definition of, 143, 144f; origins of concept, 10, 143; spatial patterns in, 126–27, 142–49

- migration: in disease ecology, 190; in metapopulations, 143, 145–48, 149f
 milkweed, 235–38, 237f
 Milne, A., 10
 mineralization, 220
 mites, 157f, 158
 monocultures, 211–12, 213f
 mortality. *See* death rate
 mosquitoes, 194–96
 Mount McKinley National Park, 73–74, 74t
 multiple cropping, 211–12, 213f
 Murdoch, W. W., 176
 Murie, O., 73, 74t
 mutualism(s), 225–38; definition of, 225; facultative, 228, 228f, 231–33, 232f, 234f; nonlinearity in, 233–35, 234f, 236f; obligate, 228–38, 228f, 231f, 232f, 234f, 236f; ubiquity of, 225
 mutualism coefficients, 226, 227f
 mycorrhizae, 235–38, 237f
Myzus persicae, 162, 162f

 native plants, demography of, 78–80, 79f
 natural selection: competition in, 198; in life history evolution, 63–69; phylogenetic constraints on, 65, 68
 natural systems agriculture, 67
 nearest-neighbor analysis, 134–35, 134f
 Nelder, J. A., 21, 22, 28–29
 net attack rate, 180–81, 181f
 neutral stability, in predator–prey interactions, 156, 158, 162, 179, 183
 niches, ecological, 209–11
 Nicholson, A. J., 9–10, 176, 179–80, 181, 182, 186
 nonlinear dynamic systems, 82, 90, 240, 242
 Noy-Meir, E., 171, 172, 174, 175

 obligate mutualism, 228–38, 228f, 231f, 232f, 234f, 236f
 olive trees, 176
 one-dimensional maps: equilibrium and stability in, 97–106; rule of projection in, 98–99, 100f; stair-stepping technique in, 98–100, 99f, 101f, 102f
 one-sided competition, 53–54
 optimal reproduction schedules, 67–72
 oscillation: eigenvalues and, 96f, 97; in models of dynamics, 86–87, 87f; in predator–prey interactions, 155–58; structural stability and, 111–12
 oscillatory point attractors: eigenvalues and, 96f, 97; in logistic map, 106–7; in one-dimensional maps, 102, 103f; physical models of, 86–87, 87f; in predator–prey interactions, 159, 162
 oscillatory point repellers: eigenvalues and, 96f, 97; in logistic map, 106–7; in one-dimensional maps, 102, 103f; physical models of, 87, 87f
Ovis dalli (Dall mountain sheep), 73–74, 74t

 Paige, K. N., 70
 panmictic populations, 142, 143, 144f
 paradoxes: of ecology, 122; in predator–prey interactions, 168–70, 176
 parakeets, 128, 128f
Paramecium bursaria, 13–14, 13f–15f, 16, 17f
 parameter space, vs. state space, 111
 parasites, in density dependence, 9, 10. *See also* disease ecology; predator–prey interactions
 patchiness, 142–43; in island biogeography theory, 144–46; and metapopulations, 142–48
Patella cochlear, 53, 53f
 pathogens. *See* disease ecology
 Pearl, R., 9
 perennial plants, 67, 72
 period-doubling bifurcation, 113, 114, 118–19, 118f
 periodic attractors: definition of, 90; physical models of, 87–89, 88f. *See also* oscillatory point attractors
 periodic repellers, physical models of, 88–89, 88f. *See also* oscillatory point repellers
 pertussis, 242
 phase plane diagrams, 87f
Pheidole protensa, 208–9
Pheidole synanthropica, 208–9
 phenotypes, in life history evolution, 66, 67–68
 phenotypic optimality models, 67–68
 phylogenetic constraints, on natural selection, 65, 68
 Pianka, E., 64
 plants: age-structured population projection matrices for, 38–39; annual vs. perennial, 67, 72; essential resources for, 219–20; invasive vs. native, demography of, 78–80, 79f; mutualism in, 232–33, 235–38, 237f; optimal reproductive schedules for, 69–70, 72; size distributions of, 51–55, 51f, 53f

- 54f; spatial distribution of, 141f; thinning laws in, 24–27, 26f, 27f, 54–55; yield–density relationship in, 17–22, 18f. *See also specific types*
- plastic effects, 24
- Poincaré section, 88–89, 88f, 89f
- point attractors: definition of, 85; eigenvalues and, 92–97, 96f; in interspecific competition, 206, 206f; in one-dimensional maps, 100, 101f; physical models of, 85, 85f, 94–95, 95f; vs. strange attractors, 90–91. *See also* oscillatory point attractors
- point pattern analysis, 134–36
- point repellers: eigenvalues and, 92–97, 96f; in interspecific competition, 204–5, 205f, 206f; in one-dimensional maps, 102, 102f; physical models of, 85, 85f; saddle, 96, 96f, 205, 207f. *See also* oscillatory point repellers
- Poisson distribution, 129–34, 181
- politics, 241–42
- Polycaryum laeve*, 187
- population(s): classes of individuals in, 30; complexity of concept, 1–2; definitions of, 1–2; true law of, 9
- population ecology: common applications of, 1; complexities in, 1–2; fields with basis in, 240–41; general principles of, xviii, 2; reasons for studying, 240–42; unit of analysis in, 1–2
- population viability analysis (PVA), 76–78
- predation: in density dependence, 9, 10; as size dependent, 51
- predator–prey interactions, 152–86; constant predator model of, 109, 113–14, 115; definition of, 152; density dependence in, 154, 158–68, 160f, 167f, 171; in discrete time, 176–86, 179f, 182f–85f; first principles of, 153–58; functional response in, 161–68, 167f, 171–76, 171f, 173f, 177f; graphical approach to, 170–76, 171f; neutral stability in, 156, 158, 162, 179, 183; oscillation in, 155–58; paradoxes in applications of, 168–70, 176
- probability: in disease ecology, 188, 194–95; in life history analysis, 68–70, 72; in predator–prey interactions, 180–81, 181f; in spatial patterns, 129–32; in stage-structured population projection matrices, 41, 45; transition, in projection matrices, 44–45, 73
- projection matrices, 30–61; applications of, 73–80; density dependence in, 48–57, 50f; eigenvalues of, 45–48, 94; elasticity of, 47–48, 75–76, 78–80; sensitivity of, 47–48, 75; transition probabilities in, 44–45, 73; as type of structured population model, 30. *See also* age-structured population projection matrices; stage-structured population projection matrices
- propagule rain, 148–49, 149f
- proportionality constant, 19, 188, 195
- Propylea dissecta*, 162, 162f
- PVA. *See* population viability analysis
- quadratic equations, 15
- Ramula, S., 78–80, 79f
- random distribution, measurement of, 129–33. *See also* Poisson distribution
- random populations, vs. chaotic populations, 10
- rape (rapeseed), 18f, 19, 19f
- reaction, in formation of spatial patterns, 137–42
- reaction/diffusion equation, 138, 140–42
- Reed, L. J., 9
- refuges, in predator–prey interactions, 174, 175f, 176
- regime shifts, 114, 115f
- regional stability, 105f, 107, 107f, 122, 123, 123f
- regression, in stage-structured projection matrices, 45
- relative growth rate (RGR), 52, 53
- relative yield total (RYT), 211–12
- remoras, 144
- repellers: in bifurcation diagrams, 116–22; definition of, 84, 84f; eigenvalues and, 92–97; in interspecific competition, 204–5, 205f, 206f; in logistic map, 106–10; in one-dimensional maps, 97–106; periodic, 88–89, 88f; physical models of, 85–90, 85f–89f; point (*See* point repellers); in predator–prey interactions, 172–74, 173f; structural stability and, 110–15, 111f; types of, 85–89; use of term, 84–85
- repressors, in spatial patterns, 142, 143f
- reproduction: cost of, 63, 65, 66–67; and density, relationship between, 11, 17–22; energy allocation to, 68–72, 69f; in life history evolution, 63–72; optimal schedules for, 67–72; as size dependent, 51
- reproductive value, 47
- rescue effect, 148–49, 149f
- resistance, disease, 190

- resource(s): essential, 219–20, 219f, 220f; substitutive, 219, 219f
- resource–consumer interactions, definition of, 152. *See also* competition; predator–prey interactions
- resource management, sustainable, stability and equilibrium in, 81
- RGR. *See* relative growth rate
- Ricker, W. E., 29
- Ricker equation, 29, 186
- Ripley's *K* analysis, 135, 135f
- r*–*K* selection, 63, 64–65, 65f
- root competition, 54, 54f
- Rose, M. R., 66, 67f
- Rosenzweig, M. L., 171, 173–74, 174f
- Ross, R., 194, 195
- RYT. *See* relative yield total
- saddle-node bifurcation, 113–14, 113f, 115f
- saddle point repellers, 96, 96f, 205, 207f
- saplings, in stage-structured projection matrices, 40–41
- scale insects, 176
- scarlet gilia (*Ipomopsis aggregata*), 70
- Schaffer, M. V., 72
- Schaffer, W. M., 68, 69, 69f, 72
- schistosomiasis, 240
- scramble competition, 53, 54
- seedlings, in stage-structured projection matrices, 40–41
- seeds, in stage-structured projection matrices, 40–41
- self-thinning laws, 11, 24–25. *See also* thinning laws
- semelparity, 69–72, 69f
- sensitive dependence on initial conditions, in chaos, 124–25
- sensitivity, of projection matrices, 47–48; applications of, 75, 78
- separatrix, 110
- sharks, 144
- sheep, 73–74, 74f
- Shinozaki, K., 18–19, 19f, 21, 22
- Shinozaki–Kira equation, 18–19, 19f, 21, 22
- shoot competition, 54, 54f
- Silvertown, J., 78
- Simberloff, D., 146
- singularities, 84. *See also* equilibrium points
- SIR model of epidemiology, 190–94, 190f, 192f, 195
- six-spotted mites (*Eotetranychus sexmaculatus*), 157f, 158
- size, of systems, stability related to, 81–82
- size-asymmetric competition, 53
- size-distributed populations: density dependence in, 50–55; mechanisms of, 51–55; projection matrices of, 40; quantification of, 51–52
- size hierarchies, 52–55
- size-symmetric competition, 53
- Slobodkin, Larry, xvii
- Smith, H. S., 9
- snowshoe hares, 156–57, 157f
- source–sink populations, 143, 144f, 147
- spatial patterns, 126–51, 241; classification of, 143, 144f; importance of scale in, 134–36; measurement of, 128–33; mechanisms of formation of, 137–42; in metapopulations, 126–27, 142–49; point pattern analysis in, 134–36; Poisson distribution in, 129–34
- species diversity, 240
- sports metaphor, 198, 205, 207
- square matrices, 37–38, 42–43, 46, 57–60
- stability, 81–125; in bifurcation diagrams, 116–22; centrality of concept, 81; complexity related to, 81–82, 122; eigenvalues for, 104, 105f; intuitive ideas of, 82, 83–91; in logistic map, 106–10, 112, 112f, 113f; neutral, 156, 158, 162, 179, 183; in one-dimensional maps, 97–106; problems with traditional theories on, 81–82, 122–23; regional, 105f, 107, 107f, 122, 123, 123f; structural, 110–15, 111f–13f, 123–24
- stable age distribution, 34–37, 35f, 45–47
- stable equilibrium points, 83–84, 84f. *See also* attractors
- stable stage distribution, 44, 45–47
- stage-structured population projection matrices, 39–45, 45f; density dependence in, 56–57; stable stage distribution in, 44, 45–47; transition probabilities in, 44–45, 73
- stair-stepping technique, 98–100, 99f, 101f, 102f
- state space: definition of, 85; for one-dimensional model, 92, 92f, 93f; vs. parameter space, 111; for two-dimensional model, 94
- state variables, 85
- static life tables, 73–74
- Stiven, A. E., 188
- strange attractors: boundaries of, 90; chaotic behavior of, 89–91, 106, 125; definition of, 89; in one-dimensional maps, 104–6, 105f; physical models of, 89–91, 89f; vs. point attractors, 90–91; structural stability and, 115

- Strong, D., 10
- structural stability, 110–15, 111f–13f, 123–24
- structured population models: age as basis for, 30–39; applications of, 73–80; definition of, 30; density dependence in, 36, 48–57; projection matrices as type of, 30 (*See also* projection matrices); stage as basis for, 39–45; transition probabilities in, 44–45, 73; vs. unstructured models, 30; uses for, 30
- subpopulations, of metapopulations, 143, 145
- substitutive resources, 219, 219f
- supply rate, 218, 220–21
- supply vectors, 220–22, 220f, 223f
- survivorship: in age-structured projection matrices, 31–32; and density, relationship between, 11; in dynamic life tables, 74–76; in life history evolution, 63–66; in stage-structured projection matrices, 40; in static life tables, 73–74
- sustainable resource management, stability and equilibrium in, 81
- Taylor, A. D., 147
- thinning laws, 22–27; self-, 11, 24–25; and size distributions, 54–55; three-halves, 25–27, 27f
- Thompson, W. R., 9
- three-halves thinning law, 25–27, 27f
- Tilman, D., 146, 147f, 216, 219
- time: in age-structured populations, 30; continuous, logistic equation in, 11; in models of dynamics, 85–87; in predator–prey interactions, 163–64. *See also* discrete time
- tipping points, 114, 115f
- tornadoes, 124–25
- transition probabilities, in projection matrices, 44–45, 73
- transmission coefficient, 188
- trees: interspecific competition among, 152; spatial distribution of, 127, 128, 131–33, 134–36; stage-structured projection matrices for, 39–45, 43f; thinning laws and, 22–24, 23f. *See also specific types*
- Tribolium castaneum*, 56–57, 119–22, 120f, 121f
- Trichechus manatus latirostris* (Florida manatee), 76–78, 77t
- Tripsacum dactyloides*, 67
- true law of population, 9
- Turing, Alan, 137, 141–42, 142f, 143f
- Turing effect, 126, 142, 143f
- two-sided competition, 54
- Typhlodromus occidentalis*, 157f, 158
- unstable equilibrium points, 83–84, 84f. *See also* repellers
- unstructured models, 30
- Vannette, R., 235
- variance-to-mean ratio, 133
- vector field: definition of, 85; in eigenvalues, 92, 92f
- Verhulst, P. F., 9
- Vochysia ferruginea*, 74–76
- Volterra, V., 153, 155, 176, 198. *See also* Lotka–Volterra equations
- von Foerster, Heinz, 1
- Wallace, Alfred Russel, 198
- Wang, I.-N., 68
- Weiner, J., 52, 54
- white oak tree, 128, 128f, 131–33, 131f, 132f, 134–35, 134f, 136, 136f
- Whitham, T. G., 70
- Wilson, E. O., 64
- witch hazel tree, 128, 128f, 131–33, 131f, 132f, 134–35, 134f, 136, 136f, 137
- yield: and density, relationship between, 11, 17–22, 19f; law of constant final, 19; maximum sustained, 17
- Yoda, K., 25, 27f
- zero net growth isocline (ZNGI), 219–22, 219f

